1 Introduction

1.1 THE INEVITABILITY OF UNCERTAINTY

Gaius Julius Caesar was assassinated on March 15, 44 BC. Do you know his last words? What is the capital of the South Asian country Bangladesh? What will be the percentage gain of the Dow Jones Industrial Average during the next six months?

Maybe the first of these questions makes you think of Shakespeare's tragedy Julius Caesar from which you might recall from Act three Scene one the widely known first half of the macaronic line 'Et tu, Brute? – Then fall, Caesar!'[1] (Shakespeare 1994, p. 1032). Caesar's last words before dying severely wounded represent a contested topic among historians. How definite is your answer? With regard to the second question, you might be fairly sure about Dhaka being the capital of Bangladesh, but you may still prefer, for whatever reason, to check your answer by consulting an up-to-date almanac. Finally, there are many relevant events to come up yet during half a year – at present almost certainly unknown to you – so that we dare to anticipate that your beliefs about the future state of the stock market will be vague to at least some degree.

Yours, ours, anybody's understanding of the past, the present and the future is naturally incomplete. It is such imperfect knowledge which implies what we commonly refer to as *uncertainty*, the natural state that inevitably attends all human activities. Throughout this book, uncertainty is regarded in a personal way, that is as a notion that describes the relationship between a person issuing a statement of uncertainty and the real world to which that statement relates – and in which

[1] A literal translation is 'And you, Brutus? Then fall, Caesar.' Alternatively, the first part of the expression may also be translated as 'You too, Brutus?' or 'Even you, Brutus?'.

Data Analysis in Forensic Science: A Bayesian Decision Perspective Franco Taroni, Silvia Bozza, Alex Biedermann, Paolo Garbolino and Colin Aitken

that person acts. A unified language for understanding and handling uncertainty, revising beliefs in the light of new information and usage for practical decision making will be the main focus of attention.

Uncertainty is an omnipresent complication in life, and the case of forensic science is no exception. As a distinct discipline of its own, along with the criminal justice system at large, forensic science is typically concerned with past events which are both unique and unreplicable. Knowledge about past occurrences is bound to be partially inaccessible, however, because of spatial and temporal limitations of our sensory capacities. Uncertainty is a fundamental problem underlying all forensic sciences. Increasingly often, it is perceived as discomforting by both scientists and other actors of the criminal justice system, which illustrates the continuing need to give it careful attention.

Notwithstanding, an objection to this attention to uncertainty is immediately possible. Suffice it to mention that past events, notably criminal activities, may generate distinct remaining traces in the form of tangible physical entities (such as blood stains or textile fibres) that can be discovered and examined. These are thought to have a potential for revealing something useful to retrace past events. The informed reader might even refer to the writings of the pioneer forensic scientist Edmond Locard, some of whose now widely quoted words are as follows:

> Nul ne peut agir avec l'intensité que suppose l'action criminelle sans laisser des marques multiples de son passage [...] tantôt le malfaiteur a laissé sur les lieux des marques de son activité, tantôt par une action inverse, il a emporté sur son corps ou sur ses vêtements les indices de son séjour ou de son geste.[2] (Locard 1920)

This quotation is still valid in these times but with the important difference that forensic scientists are now in the privileged position to analyze crime-related material of many more different kinds of nature as well as in much smaller quantities than was possible at the time of Locard's writing. A primary reason for this is that, owing to vast developments made in science and technology, today's forensic scientists have a broad scope of methods and techniques at their disposal.

Although this instrumental support offers vast capacities for providing scientists with valuable information, the outset is essentially paradoxical. On the one hand, systematic analytical testing and observation may lead to abundant quantities of information, whereas, on the other hand, such information will often be substantively lacking in the qualities that would be needed to entail (or make necessary) particular propositions[3] that are maintained by a reasoner. To state this

[2] 'No one can proceed with the intensity that the criminal act requires without leaving multiple marks of his passing [. . .] either the wrongdoer has left on the scene marks of his activity, or, on the other hand – by an inverse process – , he has taken away with him on his person (body) or clothes, indications of where he has been or what he has done.' Free translation by the authors.

[3] A proposition is interpreted here as an assertion or a statement that such-and-such is the case (e.g. an outcome or a state of nature of the kind 'the suspect is the source of the crime-stain') and also as a linguistic description of a decision. An important basis for much of the argument developed throughout this book is that it is assumed permissible to assign personal degrees of belief to propositions.

differently, forensic science is not, primarily, about the *mise en œuvre* of machinery and equipment in artifical and controlled laboratory settings. It is the general framework of circumstances within which testing is performed that makes forensic science a very challenging undertaking. For example, material related to crimes and to many real-world events in general may often be affected by contamination. The material may also be degraded and/or of such low quantity that only a single method or measurement can be applied. Sometimes it may not even be known whether the material submitted for analysis is relevant[4]. So, laboratory performance is undoubtedly important, but it is not a general guard against the rise of uncertainty in attempts to reconstruct the past.

In this book we will repeatedly come upon observations, measurements or counts, sometimes also referred to as 'raw data', and consider how such information should – in the light of uncertainty – be used to inform beliefs and support decision making. There is a practical necessity for this because such considerations represent vital steps in guaranteeing that scientific evidence meaningfully serves the purpose of a particular forensic branch of application. In order to comply with this requirement, forensic science needs to enquire about ways that allow one to learn about past events in the light of uncertain information, preferably in a manner that is in some sense rational and internally consistent.

More detailed explanation of what we mean by 'rational' and 'consistent' is delayed to a discussion in Chapter 2. For the time being, we solely note that these ways are intended to provide assistance in examining whether people's opinions about unobserved matters of fact are justified and whether these people actually have the reasons to believe that their decisions in the light of these opinions are optimal.

The very idea of enquiring about how one ought to reason and act sensibly under uncertainty takes such a central role that Ian Evett relied on it for the purpose of providing a definition of forensic science:

> [. . .] I will settle for a simple premise: forensic *science* is a state of mind, I mean that whether a particular individual is behaving, at a given juncture, as a scientist can be determined by the mental process underlying his/her actions and words. (Evett 1996, p. 121)

1.2 DESIDERATA IN EVIDENTIAL ASSESSMENT

Prior to providing a more formal introduction to the methods we seek to implement throughout this book, it is useful to set forth some very general, practical precepts to which we wish our analytic thought and behaviour to conform. Such precepts will be helpful, for example, to clarify why we will be giving preference to some methods and views rather than to others.

[4]In forensic science, 'relevance' is commonly used as a qualifier for material that has a true connection with the offence or the offender (Stoney 1994).

Consideration is given hereafter essentially to six desiderata upon which the majority of current scientific and legal literature and practice converge in their opinion. These desirable properties are balance, transparency, robustness, added value, flexibility and logic. These notions have been advocated and contextualized, to a great extent, by some quarters in forensic science and from jurists from the so-called 'New Evidence Scholarship' (Lempert 1988).

For an inferential process to be *balanced*, or in the words of some authors, impartial (Jackson 2000), attention cannot be restricted to only one side of an argument. Evett (1996, p. 122) has noted, for instance, that 'a scientist cannot speculate about the truth of a proposition without considering at least one alternative proposition. Indeed, an interpretation is without meaning unless the scientist clearly states the alternatives he has considered.' The requirement of considering alternative propositions is a general one that equally applies in many instances of daily life (Lindley 1985), but in legal contexts, its role is fundamental. Evett, during an interview, has expressed this as follows:

> Balance means that when I am doing anything for a court of law, I do it in full knowledge that there are two sides represented in that court. Even though the evidence that I've found appears to favour one or the other of those sides, my view of that evidence is directed not to proving the case for that side, but to helping the court to set that evidence and the views of both teams, prosecution and defence. (Joyce 2005, p. 37)

Note that there is more in this quotation than a sole requirement of considering alternatives. It also states that forensic scientists should primarily be concerned with the evidence and not with the competing propositions that are forwarded to explain it. This distinction is crucial in that it provides for a sharp demarcation of the boundaries of the expert's and the court's areas of competence. Failures in recognizing that distinction are at the heart of pitfalls of intuition that have caused – and continue to cause – much discussion throughout judicial literature and practice.

Besides balance, a forensic scientist's evaluation should also comply with the requirements of:

- *transparency*, that is, in the words of Jackson (2000, p. 84), '[. . .] explaining in a clear and explicit way what we have done, why we have done it and how we have arrived at our conclusions. We need to expose the reasoning, the rationale, behind our work.'
- *robustness*, which challenges a scientist's ability to explain the grounds for his opinion together with his degree of understanding of the particular evidence type (Jackson 2000).
- *added value*, a descriptor of a forensic deliverable that contributes in some substantial way to a case. Often, added value is a function of time and monetary resources, deployed in a way such as to help solve or clarify specific issues that actually matter with respect to a given client's objectives.

These desiderata characterize primarily the scientist, that is his attitude in evaluating and offering evidence, as well as the product of that activity. The degree to which the scientist succeeds in meeting these criteria depends crucially on the chosen inferential framework, which may be judged by the following two criteria:

- *flexibility*, a criterion that demands a form of reasoning to be generally applicable, that is, not limited to particular subject matter (Robertson and Vignaux 1998).
- *logic*, that is, broadly speaking, a set of principles that qualify as 'rational'. In turn, that rational system must also conform, as will be explained later in Chapter 2, to certain minimum requirements (Robertson and Vignaux 1993).

These last two issues – properties of an inferential method rather than behavioural aspects of the scientist – represent the principal topics to which the subsequent parts of this book thematically connect.

1.3 THE IMPORTANCE OF THE PROPOSITIONAL FRAMEWORK AND THE NATURE OF EVIDENTIAL ASSESSMENT

A few additional remarks are necessary on the requirement of balance, a criterion described so far as one that requires a scientist to consider at least two competing propositions.

First, attention should be drawn to the exact phrasing of propositions, an idea that underlies a concept known in the context as propositional level or hierarchy of propositions (Cook *et al*. 1998). The reasons for this are twofold. On the one hand, a proposition's content crucially affects the degree to which that proposition is helpful for the courts. For example, the pair of propositions 'the suspect (some other person) is the source of the crime stain' (known in the context as a source-level proposition) addresses a potential link between an item of evidence and an individual (that is, a suspect) on a rather general level. Generally, activity-level (e.g. 'the suspect (some other person) attacked the victim') or crime-level (e.g. 'the suspect (some other person) is the offender') propositions tend to meet a court's need more closely. On the other hand, the propositional level defines the extent of circumstantial information that is needed to address a proposition meaningfully. For example, when reasoning from a source- to a crime-level proposition, consideration needs to be given to the relevance of a crime stain (that is, whether or not it has been left by the offender), an aspect that is not necessarily needed when attention is confined to a source-level proposition.

Secondly, given a proposition of interest, forensic scientists usually assess the relative degrees to which evidence is compatible with the various settings (that is, the propositions) under consideration. The question, however, of what the believability of each setting actually ought to be, is not an issue for forensic scientists.

Addressing a target proposition requires – for reasons given later in Chapter 2 – a belief-state prior to the consideration of new facts as well as profound knowledge of circumstantial information. Forensic scientists cannot comply with any of these requirements. Even if they could, their focus on an issue (e.g. a proposition of the kind 'the suspect is the source of the crime stain') rather than on the evidence would amount to usurping the role of the court (Aitken and Taroni 2004).

Thirdly, it is worth insisting on having a well-defined framework of propositions. This is in sharp contrast to occasionally held opinions according to which data should be allowed to 'speak for themselves', a suggestion that evidential value represents some sort of intrinsic attribute. This is viewed cautiously in forensic science, where the following position has been reached:

> In court as elsewhere, the data cannot 'speak for itself'. It has to be interpreted in the light of the competing hypotheses put forward and against a background of knowledge and experience about the world. (Robertson and Vignaux 1993, p. 470)

As may be seen, the concept of propositions is important because it is closely tied to the notion of evidential value. For the time being, we tentatively consider the latter as a personalistic function of the former, in the sense that value is assigned to evidence by a particular individual depending on the propositions among which that individual seeks to discriminate and auxilary contextual information that is available to that individual. Arguably, evidential value is neither seen as an abstract property of the external world nor as one that can be elicited in a uniquely defined way.

Generally, the propositional framework is organized as part of an evaluative procedure, that is a model that specifies the relevant ingredients of an inferential process, their relationships along with rules that state how these elements ought to be used for inference. The issue that relates this brief mention of the propositional framework to the main topic of this book – data analysis – is the fact that the latter is needed in order to enable the former to provide quantitative expressions that are appropriate for the purpose for which a particular inference procedure has been designed.

1.4 FROM DESIDERATA TO APPLICATIONS

Although the general criteria to which we would like evidential assessment to conform may appear intuitively reasonable, it may be far from obvious how to implement them to bridge practical difficulties associated with forensic science as a discipline of reasoning and acting under uncertainty. As for themselves, the stated principles describe desirable, abstract properties rather than explicit ways in which one ought to proceed. The criteria – if met by the scientist – should contribute to the avoidance of the reduction of forensic expertise to ad hoc guesswork and unwarranted claims of 'many years of experience' (Evett 1996). Beyond this, however, the mere statement of the principles also remains insufficient for the need.

Further concepts and discussion are thus needed for examining whether scientists' analyses, evaluations and reportings are trustworthy. Among these is an approach to the description of uncertainty as well as rules that prescribe the combination of expressions of uncertainty. For this purpose, Chapter 2 will outline in detail a method for calculating with beliefs that is part of a package that also contains a procedure to use personal beliefs to inform decision making. As will be seen, these elements represent the fundamental tenets of the discipline of *statistics* (Lindley 2000b).

In this book we argue that statistics is a beneficial resource with important capacities for both clarifying and analyzing a wide range of practical problems. In particular, we will focus on statistical concepts that allow one to make plain and conceptualize the passage from the assessment of uncertainties associated with evidence to the assessment of uncertainties associated with particular explanatory propositions, including consistent choices amongst them. We justify this focus of enquiry by our conviction that these concepts have a substantial potential for the enhancement of the quality of forensic expertise.

Although we will argue that the methods yet to be introduced in later sections and chapters are the most appropriate ones currently available, we will not address the separate issue of how evidence is best presented before trial. This latter issue is a distinct topic of its own which, in the context, is also referred to as the 'presentation problem' (Redmayne 2001). This topic extends to additional complications that touch on discussions beyond the scope of this book. Evett and Weir (1998, p. 29) expressed this point concisely when they wrote that, 'in particular, we are going to take the evidence into the court room, where the proceedings owe no allegiance to the laws of science or mathematics and many of the participants are stubbornly nonnumerate.'

Notwithstanding the above, this book's central points of attention – forensic inference and decision analysis – draw their legitimacy from cogent practical reasons. An illustrative example for this is provided by courts that, typically, seek to reduce their uncertainty about a defendant's true connection with a criminal act (Lindley 2006). Often, part of this effort is thought to be achieved on the basis of evidence as offered by forensic scientists. According to this view, evidential assessment, that is a process of reasoning under uncertainty, constitutes a preliminary to judicial decision making (e.g. deciding if a suspect should be found guilty for the offence for which he has been charged) and taking such assessment seriously reflects the intention of promoting accurate decision making (Fienberg and Schervish 1986; Kaplan 1968; Kaye 1988; Redmayne 2001; Robertson and Vignaux 1993).

In particular, a decision-based approach can help (i) to clarify the fundamental differences between the value of evidence as reported by an expert and the final decision that is to be reached by a customer, and (ii) to provide a means to show a way ahead as to how these two distinct roles can be conceptualized to interface neatly with each other. Both of these are topics that are currently viewed differently rather than in a unified manner. This illustrates the continuing need for research in this area.

1.5 THE BAYESIAN CORE OF FORENSIC SCIENCE

Prior to proceeding with more technical chapters we anticipate at this point – still with the intention of relying on an essentially informal style of presentation – some of the main arguments and topics that will be advocated throughout this book, while reserving elements of logical and philosophical justifications to later discussion.

One of the credences of which we seek to convince the reader is that uncertainties about propositions should be expressed by the concept of probability. To this viewpoint we immediately add, however, that we will be giving preference to probability theory employed in that of its distinct interpretations which views probabilities as degrees of belief, a standpoint commonly known as the subjectivist (or personalist) interpretation of probability. Such degrees of belief are personalized assessments of credibility formed by an individual about something uncertain, given information which the individual knows or can discover. In short, the probability apparatus will be used as a concept of reference to which personalized weights to the possible states of the uncertain world that surrounds us may be attached.

Even though such uncertainty is inevitable, recall from Section 1.1 that we live in a world in which further information may be gained by enquiry, analysis and experimentation. As a consequence of this, some means is required to adjust existing beliefs in the light of new evidence. A second credence which will thus be emphasized here is that the revision of beliefs should be operated according to Bayesian procedures. The term ‘Bayesian’ stems from a theorem – Bayes’ theorem (Section 2.3.1) – that is a logical consequence of the basic rules of probability. We will repeatedly come across the theorem because it is a very important result that helps one to understand how to treat new evidence. As an aside, we note that, although the theorem has about a 250-year history, the attribute ‘Bayesian’ as a descriptor of a particular class of inference methods appears to have gained more widespread use only since the middle of the twentieth century.

Given a set of beliefs about the unknowable states of the world, the general objective is to identify an available course of action that is logically consistent with a person’s personal preferences for consequences. This is an expression of a view according to which one decides on the basis of essentially two ingredients. These are, on the one hand, one’s beliefs about past, present or future happenings and, on the other hand, one’s valuation of consequences. As noted above, the former will be expressed by probability. The latter will be captured by invoking an additional concept, known as utility. Both concepts can operate within a general theory of decision that involves the practical rule which says that one should select that decision which has the highest expected utility (or, alternatively, which minimizes expected loss). When the class of such operations is based on beliefs that have received a Bayesian updating (statistical inference), then this process is called Bayesian decision analysis.

Both within and outside forensic science, the Bayesian package for inference and decision is considered – with continually increasing agreement – as the currently most appropriate and comprehensive approach to the various issues pertaining to the assessment of scientific evidence. In a legal context, the concept is particularly relevant because of the support it provides in conforming with the principles and requirements set forth in Section 1.2. Subscription to the Bayesian decision approach, however, does not suggest that the approach is perfect, a point that is noted by Evett and Weir (1998, p. 29):

> It is not our claim that Bayesian inference is a panacea for all problems of the legal process. However, we do maintain that it is the best available model for understanding the interpretation of scientific evidence.

Practical applications of patterns of reasoning corresponding to a Bayesian approach can be found, for example, as early as the beginning of the 20th century (Taroni *et al*. 1998). Bayesian ideas for inference then entered legal literature and debates more systematically only in the second half of the twentieth century. Kingston (1965a), Finkelstein and Fairley (1970) and Lindley (1977a) are some of the main reference publications from that period. Later, within the 1990s, specialized textbooks from Aitken and Stoney (1991), Aitken (1995) and Robertson and Vignaux (1995) appeared. During the past decade, further textbooks – along with a regular stream of research papers – focusing on Bayesian evaluations of particular categories of evidence, such as glass (Curran *et al*. 2000) or DNA (Balding 2005; Buckleton *et al*. 2004) were published. Compared to this, decision analysis is a rather sparsely studied area, in particular within forensic science. Some of the few available references are mostly from legal scholars, mentioned earlier in Section 1.4. However, for forensic science applications, decision making (e.g. about target propositions of interest) is a presently latent topic with room for many thought-provoking and interesting issues that ask for explanations and the formulation of effective answers.

It is worth noting that many discussions of probabilistic reasoning applied to forensic and legal matters in general rely on probability as a concept that is defined on propositions, that is linguistic entities of the kind that were presented at the beginning of Section 1.1. A probability statement about such entities – that is to say, some form of proposition or hypothesis – may typically have that character of a personalized expression of belief, as was mentioned earlier in this section.

Besides this, there is an additional facet of probability relevant to forensic applications, known as the set-theoretical development of probability, introduced in the last century by the Soviet mathematician A. N. Kolmogorov (Kolmogorov 1933). In that development, probability is defined on subsets of some given set. It is customary to denote the latter set as one that comprises all elementary possibilities, often referred to as outcomes of an experiment (sample space). This approach lends itself to a series of extensions that makes it of particular interest for the fields of mathematics and statistics (e.g. due to the applicability of the full differential and integral calculus).

The latter development of probability is used essentially in contexts where the main issue is uncertainty about the true value of a parameter[5], such as a mean or a proportion. The aim then is to use the probability calculus to obtain probability statements about such parameters. To return to the above-mentioned argument, these probability statements have a personalized interpretation in terms of degrees of belief and such degrees are revised by the use of Bayes' theorem, operating after the provision of data. Since one is concerned with a set of objects or individuals (a population), one can usually extract a subset of that set and investigate it. The result of this is a set of numbers – the data – for which a statistic (e.g. a mean) may be calculated and used as a basis for revising beliefs about the population parameter held prior to inspection of those data. When in addition to that, the goal is to choose a particular number as an estimate for the parameter, then this can be conceptualized as a decision problem using the ingredients informally introduced above, that is decisions and utilities.

These are, in brief, the main aspects of the Bayesian approach to statistics and decision analysis, a more formal and detailed account of which is given in the next two chapters.

1.6 STRUCTURE OF THE BOOK

In the previous sections of this chapter, a discussion was initiated on Bayes' theorem, presented essentially as a formalization of logic and common sense which makes it a valuable tool for reasoning about evidence in situations involving uncertainty. In fact, forensic practice routinely involves the collection of sets of observations or measurements, but they may be compatible with several distinct hypotheses of interest given and some of them are (a priori) less plausible than others. The problem, as stated, then consists in drawing conclusions in such situations while the idea underlying the inference analysis consists in offering guidance and influencing one's behaviour (Cornfield 1967).

Bayes' theorem plays a central role in quantifying the uncertainty associated with particular conclusions. The forthcoming Chapters, 2 and 3, provide a careful outline of this role and its relation to the process of consistent choice between hypotheses when available evidence is imperfect. The viewpoint will be that of a unification of the theories of (subjective) probabilities (mainly Bayes' theorem) and utility within decision theory in order to set forth the construction of a co-ordinated and structured whole.

Combining both theoretical elements and practical examples, Part II of the book – Chapters 4 to 8 – will proceed with focusing in more detail on the idea

[5]For the time being, a parameter is taken as a characteristic of the distribution of the measurements on, or categorization of, the entire set of members (e.g. individuals or objects) of a target population.

of 'judgements', that is, stated otherwise, assessments, considered decisions or sensible conclusions (not, however, in the sense of decisions of a court or judge). Actually, day-to-day forensic practice involves judgements, in a variety of facets and in contexts. For the purpose of illustration, consider the following:

1. How is one to judge an estimate of, say, the proportion of red car paint flakes found on a victim of a car accident in a given town, or the alcohol concentration given a series of measurements on a blood sample?
2. How is one to assess whether the value of a parameter of interest, such as the colour dye concentration in ecstasy tablets, lies within a given interval?
3. How is one to decide among competing hypotheses according to which, for example, a rate is greater (or not) than a given value, or two series of continuous measurements differ (or do not differ)?
4. How ought the value of a particular item of evidence be assessed?
5. How many samples should be analyzed in a consignment of, say ecstasy pills seized in a criminal investigation?
6. How ought one to proceed when an unknown item needs to be associated with or arranged into one of different specific (and known) classes? In other words, how is one to judge the appropriateness of two (or more) competing models for a given forensic real-world phenomenon?

Questions of this kind will be described, exemplified, analyzed and commented on in the light of Bayesian statistical methodology. Specifically, Chapter 4 approaches problems in point estimation, as given by question 1 above. Credible intervals, addressed by question 2, are treated in Chapter 5. Chapter 6 focuses on hypothesis testing to approach issues to which questions 3 and 4 relate. Question 5 relates to sampling problems, a recurrent topic in many forensic science disciplines, and considered in Chapter 7. Finally, questions such as question 6 are studied in Chapter 8.

The book focuses on a versatile list of statistical questions that may reasonably be encountered in forensic practice. The aim is to show how the Bayesian framework can be developed, understood and practically applied. Although the Bayesian approach has been in place for some time already and it may be tempting to think of it as an 'old tool', actual practice demonstrates that its help is more than ever indispensable for addressing current problems in evaluating forensic evidence. This, then, is an instance where we seek to answer a viewpoint taken by Good (Good 1962, p. 383), according to which it may be beneficial to 'learn to use a little of the language of the theory of rationality to understand what it means to be reasonably consistent. "Rational" is to be interpreted in relation to the theory of rationality, namely Bayes' theory, in which emphasis is on judgments of probabilities, utilities and expected utilities.'