AUTHORS' PREPRINT

Published as: Taroni F., Biedermann A., Aitken C., Statistical Interpretation of Evidence: Bayesian Analysis, Encyclopedia of Forensic Sciences (Third Edition), Max M. Houck (Ed.), Oxford: Elsevier, 2023, Pages 656–663, ISBN 9780128236789, https://doi.org/10.1016/B978-0-12-823677-2.00171-9. (https://www.sciencedirect.com/science/article/pii/B9780128236772001719)

<u>Note:</u> *This is an uncorrected preprint. For reference purposes, readers should rely on the final published version available at https://www.sciencedirect.com*

Statistical Interpretation of Evidence: Bayesian Analysis

Franco Taroni,
University of Lausanne, Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, Lausanne (Switzerland), franco.taroni@unil.ch

Alex Biedermann,
University of Lausanne, Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, Lausanne (Switzerland), alex.biedermann@unil.ch

Colin Aitken
University of Edinburgh, School of Mathematics, Edinburgh, UK

Abstract

Probability theory provides the general framework within which assignments of probabilities of past, present, and future events are coherently modified in the light of observed events or, more generally, new information. Forensic scientists, as an illustrative example, routinely face tasks of reasoning under uncertainty when they seek to assist members of the judiciary in evaluating or interpreting the meaning of items of scientific evidence. As a consequence of the laws of probability theory and related concepts, Bayes' theorem is the key rule according to which to conduct such reasoning in order to comply with the requirement of rationality. This quantification, though, does not represent the end of the matter as the forensic scientist may also be confronted with questions of how to make a rational choice amongst alternative courses of action. This article presents the role of Bayes' theorem, and its extension to decision analysis, in categorical and continuous data analysis in forensic science applications. It emphasizes the importance of propositional hierarchies, the role of background information, the interpretation of probability as personal degrees of belief and the personal quantification of the consequences of decisions. The discussion also includes a sketch of some common pitfalls of intuition associated with probabilistic reasoning in legal contexts.

Keywords

Bayes' factor, Bayes' theorem, Categorical data, Continuous data, Decision theory, Degree of belief, Evidence evaluation, Fallacy, Interpretation, Likelihood ratio, Posterior probability, Prior probability, Probability theory, Subjective probability, Utility.

Key points

- Probability is the reference framework for scientific reasoning under uncertainty
- Bayes' theorem, a consequence of probability theory and related concepts, provides a way to formalise the assessment of the value of the evidence with respect to propositions of judicial interest
- Decision theory, an extension of probability theory, can shed light on how to use scientific evidence (in the form of discrete and continuous data) for coherent decision analysis under uncertainty

Introduction

Bruno de Finetti, a pioneering subjective probabilist, considered that the role of probability theory in inductive logic is to show how the evaluations of probabilities of unobserved events are to be modified in the light of observed events, and that this translates, in the mathematical formulation of induction, the meaning of the phrase 'to learn from experience'. Forensic scientists, as an illustrative example, routinely face inductive reasoning when they ask to assist in evaluation or interpretation the meaning of items of scientific evidence. This directs attention to Bayes' theorem, which, in essence, formalizes induction.
In Bayesian analysis, all available information is used in order to reduce the extent of uncertainty associated with an inferential problem. As new information is obtained, it is combined with any previous information, and this forms the basis for statistical procedures. The formal mechanism used to combine new information with previously available information is generally known as Bayes' theorem. Bayes' theorem involves the use of probabilities because probability can be thought of as the coherent language of uncertainty. At any given point in time, the scientist's state of information about some uncertain event (or quantity) can be represented by a set of probabilities. When new information is obtained, these probabilities are revised so as to represent all the available information. The idea of 'revising' probabilities, though, is not one that should be interpreted as a 'correction'. An updated probability is not a correction or a better evaluation of an initial probability, but solely a different probability, because it is conditioned by a new (extended) state of information.
The statistical evaluation and interpretation of evidence thus relies on a rule that specifies the dependencies amongst uncertain events through conditional probabilities. This rule enables one to define the value of evidence in the sense of the effect that evidence has on beliefs in an issue, such as whether the person of interest is the person who has committed a specific crime. The underlying ideas can be applied to categorical and continuous data. They can also be applied to situations in which there are no, or limited, data but in which there are personal assessments. They are used to ensure a coherent structure in the evaluation of items of evidence.
The Bayesian paradigm can be extended to encompass the perspective of decision making. A decision problem arises when there is a list of possible courses of action and there is uncertainty regarding the consequences that may arise from each of these courses. Bayesian decision analysis offers a transparent framework for a rational and coherent justification of one's choices.

Bayes' Theorem

The Bayesian approach is named after the Reverend Thomas Bayes, a nonconformist preacher of the eighteenth century. To him is attributed an important result that shows how uncertainty about an event, say R, can be changed by the knowledge of another event, say S:

$$\Pr(R|S) = \Pr(S|R)\Pr(R)/\Pr(S),$$

where Pr denotes probability and the vertical bar | denotes the conditioning. Thus, $\Pr(R|S)$ is the probability that R occurs, given that R has occurred. Probabilities are values between 0 and 1. The value 0 corresponds to an event that is impossible to happen, and 1 to an event that is certain to happen. Probabilities are most appropriately interpreted as subjective – in the sense of 'personal' – expressions of degrees of belief held by an individual. As such they reflect the degree of imperfection of an individual's knowledge. Such belief is graduated: as evidence accumulates, one can believe in the truth of an event more or less than one did before, one can believe more in the truth of one event than in the truth of another event, etc. The fundamental principle here is that the degrees of belief of a rational individual obey the rules of probability. Probability thus represents the quantified judgment of a particular individual. Because a probability is a measure of degree of belief rather than a long-run frequency (as suggested by other interpretations of probability), it is perfectly feasible to assign a probability to an event that involves a nonrepetitive situation. This makes the interpretation of probability, based on measures of belief, particularly useful for judicial contexts.

An alternative version of Bayes' theorem (sometimes known as Bayes' rule) is its odds form, where $\bar{R}$ denotes the complement of R so that $\Pr(\bar{R}) = 1 - \Pr(R)$. Then the odds in favor of R are $\Pr(R)/\Pr(\bar{R})$, denoted $\mathrm{O}(R)$, and the odds in favor of R given that S has occurred are denoted $\mathrm{O}(R|S)$. The odds form of Bayes' theorem is then:

$$\mathrm{O}(R|S) = \frac{\Pr(S|R)}{\Pr(S|\bar{R})} \times \mathrm{O}(R).$$

In forensic science, S, R and $\bar{R}$ are generally replaced by E, H_p, and H_d, where E is the scientific evidence, H_p is the proposition[1] proposed by the prosecution, and H_d is the proposition proposed by the defence. Thus, one has:

$$\mathrm{O}(H_p|E) = \frac{\Pr(E|H_p)}{\Pr(E|H_d)} \times \mathrm{O}(H_p).$$

The left-hand side of the equation is the odds in favor of the prosecution proposition after the scientific evidence has been presented. This is known as the *posterior odds*. The odds $\mathrm{O}(H_p)$ are the prior odds (i.e., odds prior to the presentation of the evidence). The factor that converts prior odds to posterior odds is the ratio

$$\frac{\Pr(E|H_p)}{\Pr(E|H_d)},$$

known as the *Bayes' factor*. In forensic contexts, it is regularly termed 'likelihood ratio' and abbreviated by V, short for 'value.' It can take values between 0 and ∞. A value greater than one provides support for the prosecution's proposition H_p (over H_d), and a value less than one favors the defence's proposition H_d (over H_p). Evidence for which the value is one is neutral in that the evidence does not discriminate between the two propositions of interest. Note that if logarithms are used, the relationship becomes additive. This has a very pleasing intuitive interpretation of weighing evidence in the scales of justice. The logarithm of the Bayes' factor is known, after the works of the statistician I. J. Good, as the 'weight of evidence'. Note also that it is not necessary for the two propositions in the terms denoted $\mathrm{O}(R)$ and $\mathrm{O}(R|S)$ above to be complementary; the rule still holds. Thus, the prosecution and defence propositions do not need to be complementary, though they need to be mutually exclusive.

The probative value of scientific evidence is assessed by determining a value for the Bayes' factor. The proper task of forensic scientists is the determination of that value. The role of judge and jury will be that of assessing the prior and posterior odds. Scientists can inform recipients of expert information on how to change their prior odds in the light of the evidence, but scientists cannot by themselves assign a value to the prior or posterior odds. In order to assign such a value, all the other evidence in a case has to be considered.

The terms 'evaluation' and 'interpretation' are sometimes considered synonyms, but it is helpful to conceive of a distinction. 'Evaluation' is the determination of a value for the Bayes' factor. 'Interpretation' refers to the meaning attached to this value.

The Value of Evidence

The evaluation of scientific evidence may be thought of as the assessment of a comparison. This comparison is between qualities (such as genetic traits) or results of measurements (such as refractive indices of glass fragments) of crime-related (recovered) material and of control (potential source) material. For the assessment of scientific evidence, it is widely accepted that the forensic scientist should consider at least a pair of competing propositions, in the context commonly denoted H_p and H_d. These propositions are formalized representations of the framework of circumstances. Their formulation is a crucial basis for a logical approach to the evaluation of evidence. A classification developed mainly by researchers in the United Kingdom during the late 1990s, referred to as a 'hierarchy of propositions,' considers three main categories or levels. It involves the so-called 'source,' the 'activity,' and the 'crime' level.

Categorical Data and Discrete Propositions

Evaluation given Source Level Propositions

The assessment at source level depends on analyses and measurements on the recovered (of unknown origin) and control (of known origin) material. The value of a trace (or a stain) under source level propositions, such as 'Mr. X's pullover is the source of the recovered fibers' and 'Mr. X's pullover is not the source of the recovered fibers' (i.e., another clothing is the source of the trace), does not need to take account of much more than the analytical information obtained during laboratory examination. The probability of the evidence under the first proposition (numerator of the Bayes' factor) is considered from a comparison between two items (the recovered and the control) assuming that they have come from the same source. The probability of the evidence under the second proposition (denominator of the Bayes' factor) is considered by comparison of the characteristics of the control

[1] At the time of the second edition of this Encyclopedia, the word 'hypothesis' was used to denote theories put forward by the prosecution and defence. Current thinking prefers the word 'proposition' as 'hypothesis' has connotations with statistical hypothesis testing. 'Proposition' is used throughout this article in place of 'hypothesis' from the second edition.

and recovered items in the context of a relevant population of alternative sources. The population from which the source may be thought to have come is called the relevant population.
Consider a case in which n textile fibers have been left at the scene of the crime by the person who committed the crime. A person, Mr. X, has been arrested and it is required to establish the strength of the evidence regarding competing source propositions. A forensic scientist compares the results of measurements of the physical/chemical characteristics of the questioned fibers and those taken from the Mr. X's pullover. The two propositions of interest are H_p, the questioned fibers come from Mr. X's pullover, and H_d, the recovered fibers come from some garment other than that of Mr. X. The evidence E has two parts: y, the characteristic, Γ, say, of the recovered fibers, and x, the characteristic Γ, say, of Mr. X's pullover. If the recovered fibers and Mr. X's pullover have different (incompatible) characteristics, then Mr. X's pullover would not be investigated in further detail.
Let I denote the background information. This could include (eyewitness) evidence concerning the type of garment worn by the criminal, for example. The value of the evidence is then

$$\frac{\Pr(E|H_p,I)}{\Pr(E|H_d,I)} = \frac{\Pr(y,x|H_p,I)}{\Pr(y,x|H_d,I)} = \frac{\Pr(y|x,H_p,I)}{\Pr(y|x,H_d,I)} \times \frac{\Pr(x|H_p,I)}{\Pr(x|H_d,I)}.$$

Consider two assumptions: (1) the characteristics of Mr. X's pullover are independent of whether his pullover is the source of the recovered fibers (H_p) or not (H_d), and thus $\Pr(x|H_p,I) = \Pr(x|H_d,I)$. (2) if Mr. X's pullover was not the source of the recovered fibers (H_d), then the evidence about the fibers at the crime scene (y) is independent of the evidence (x) about the characteristics of Mr. X's pullover, and thus $\Pr(y|x,H_d,I) = \Pr(y|H_d,I)$.
Hence

$$\frac{\Pr(y|x,H_p,I)}{\Pr(y|H_d,I)}.$$

The scientist knows, in addition, from data previously collected (population studies) that fiber type Γ occurs in $100\gamma\%$ of some relevant population, say Ψ.
Assuming that Mr. X's pullover is the source of the recovered fibers, the probability that the recovered fibers are of characteristic Γ, given Mr. X's pullover is of characteristic Γ, is 1. Thus, the numerator of V is 1. Alternatively, it is assumed that Mr. X's pullover was not the source of the recovered fibers. The relevant population is deemed to be Ψ. The true donor of the recovered fibers is an unknown member of Ψ. Evidence y is to the effect that the crime fibers are of characteristic Γ. This is to say that an unknown member of Ψ is Γ. The probability of this is the probability that a fiber donor drawn at random from Ψ has characteristic Γ, which is γ. Thus

$$V = \frac{1}{\gamma}.$$

This value, $1/\gamma$, is the value of the evidence of the characteristics of the recovered fibers when the source garment belongs to a member of Ψ. Given that γ is a value between 0 and 1, the Bayes' factor is greater than 1, and the evidence is said to have the value $1/\gamma$. That is, it is $1/\gamma$ times more probable to observe this evidence if Mr. X's pullover was the source of the recovered fibers than if it were not. Qualitative (i.e., verbal) scales have been proposed and they are intended to make it easier to convey the meaning of the numerical value of the evidence. However, there is ongoing discussion about the degree to which this aim has been achieved.

Evaluation given Activity Level Propositions

This hierarchical level relates to an activity. It requires that the definition of the propositions of interest include an action. Such propositions could be for example, 'Mr. X sat on the car driver's seat at the relevant time,' and 'Mr. X never sat on the car driver's seat.' The consequence of this activity – the sitting on a driver's seat – is a contact between the driver and the seat of the car. Consequently, a transfer of material (i.e., fibers in this example) may be expected. Therefore, the scientist needs to consider more detailed information about the case under examination. This information will help the scientist assess aspects such as the transfer and persistence of fibers on the driver's seat. This shows that findings cannot be evaluated given activity level propositions without a framework of circumstances.
Consider, for the sake of illustration, the following case. A crime has been committed during which the blood of a victim has been shed. A person of interest (POI) has been arrested. A single blood stain of genotype Γ has been found on an item of the POI's clothing. The POI's genotype is not Γ. The victim's genotype is Γ. There are two possibilities: T_0: the blood stain came from some background source; T_1: the blood stain was transferred during the commission of the crime.
As before, there are two propositions to consider, H_p: the POI assaulted the victim, and H_p: an unknown person assaulted the victim (i.e., the POI is not involved in any way whatsoever with the victim).

The evidence E to be considered is that a single blood stain has been found on the POI's clothing and that it is of genotype Γ. The information that the victim's genotype is Γ is considered as part of the relevant background information I. A general expression of the value of the evidence then is $V = \Pr(E|H_p, I)/\Pr(E|H_d, I)$.

Consider the numerator first and event T_0 initially. This supposes that a 'contact' of a certain intensity and duration occurred between the POI and the victim, but no blood transfer to the POI. This is an event with probability $\Pr(T_0|H_p, I)$. Also, a stain of genotype Γ must have been transferred by some other means, an event with probability $\Pr(B|\Gamma)$ where B refers to the event of a transfer of a stain from a source (i.e., a background source) other than the crime scene (here the victim).

Next, consider T_1, the event of blood transfer to the POI, an event with probability $\Pr(T_1|H_p, I)$. Given T_1, H_p and the genotype Γ of the victim, it is certain that the transferred stain is Γ. This assumes also that no blood has been transferred from a background source.
Let $t_0 = \Pr(T_0|H_p, I)$ and $t_1 = \Pr(T_1|H_p, I)$ denote the probabilities of no stain or one stain being transferred during the course of the crime. Let b_0 and b_1, respectively, denote the probabilities that a person from the relevant population will have zero blood stains or one blood stain on clothing. Let γ denote the probability that a stain acquired innocently on the clothing of a person from the relevant population will be of genotype Γ. This probability may be different from the proportion of individuals in the general population which are of type Γ. Then $\Pr(B|\Gamma) = \gamma b_1$ and the numerator can be written as $t_0\gamma b_1 + t_1 b_0$. This expresses that the presence of a stain of type Γ depends on the probability of there being no transfer (t_0), times there being such a stain as background (γb_1), plus the probability of transfer of such a stain (t_1), times the probability of there being no such stain beforehand (b_0).

Now consider the denominator where it is assumed that no physical interaction occurred between the POI and the victim. The presence of the stain is then present by chance alone. The denominator thus takes the value $\Pr(B|\Gamma)$, which equals γb_1. In summary, the value of the evidence is thus

$$V = \frac{t_0\gamma b_1 + t_1 b_0}{\gamma b_1}.$$

Extensions to cases involving transfer in the other direction (from perpetrator to scene/victim rather than from scene/victim to perpetrator), for example, or generalizations involving n stains and k groups are available in the specialized literature on the topic.

Evaluation given Crime Level Propositions

At the 'crime level' (also known as the 'offence level'), propositions are closest to those of interest to the jury. A formal development of the likelihood ratio under 'crime level' propositions shows that two additional parameters are of interest: (1) one concerns material that may be 'relevant,' meaning that it came from the offender (it is relevant to the consideration of the POI as a possible offender), (2) the other concerns the recognition that if the material is not relevant to the case, then it may have arrived at the scene from the POI for innocent reasons. As for activity level propositions, discussion of transfer in the other direction (i.e. scene/victim to perpetrator) is available in the specialized literature on the topic.
Consider the following two propositions of interest: H_p: the POI is the offender, and H_d: an unknown person is the offender.
Notice the difference between these propositions and those of the previous sections on source or activity level. At source level, the propositions referred to the POI being, or not being, the donor of the recovered trace found at the crime scene. Now, the propositions are stronger, because they specify the POI as a possible offender.
In the formal development of the likelihood ratio, a connection is needed between what is observed (i.e., the stain at the crime scene) and the propositions according to which the POI is or is not the offender. The connection is made in two steps. The first is the consideration of a proposition that the crime stain came from the offender and the alternative proposition that the crime stain did not come from the offender. If it is assumed that the crime stain came from the offender, the second step is the consideration of a proposition that the crime stain came from the POI and the alternative that the crime stain did not come from the POI.
Developing the likelihood ratio in view of these two pairs of propositions introduces the concepts of (1) 'relevance probability,' usually denoted r, and (2) 'innocent acquisition probability,' usually denoted as a. The resulting expression of the value of the evidence takes the following form:

$$V = \frac{r + \gamma'(1-r)}{\gamma r[a + (1-a)\gamma'](1-r)}.$$

Note the difference between two terms γ and γ', expressing the rarity of the corresponding characteristic. In fact, γ' is the probability that the crime stain would be of a given type, if it had been left by an unknown person who was unconnected with the crime. The population of people who may have left the stain is not necessarily the same as the population from which the criminal is assumed to have come. For DNA evidence, however, it may be acceptable to assume $\gamma = \gamma'$.

Continuous Data and Discrete Propositions

A seminal paper in 1977 by Dennis Lindley showed how the Bayes' factor could be used to evaluate evidence given by continuous data in the form of measurements. The measurements used by Lindley by way of illustration were those of the refractive index of glass. There were two sources of variation in such measurements, the variation within a window and the variation between different windows. Lindley showed how these two sources of variation could be accounted for in a single statistic. He was also able to account for the two factors which are of importance to a forensic scientist: (1) the similarity between the recovered and control materials (or items) and (2) the typicality of any perceived similarity. When the data are in the form of continuous measurements, the Bayes' factor is a ratio of probability density functions rather than a ratio of probabilities.

Consider a set $\boldsymbol{x}$ of measurements on control material and another set $\boldsymbol{y}$ of measurements on recovered material. The measurements refer to a particular characteristic, such as the refractive index of glass. For this example, $\boldsymbol{x}$ would be a set of measurements of refractive indices on fragments of a broken window at the crime scene and $\boldsymbol{y}$ a set of measurements of refractive indices on fragments of glass found on a POI. If the POI was the person who broke the window at the crime scene, the expectation is to find fragments whose refractive index corresponds to the broken window at the crime scene. If the POI was not at the crime scene, then the fragments have come from some other, unknown, source.
The quantitative part of the evidence concerning the glass fragments in this case can be denoted by $E = (\boldsymbol{x}, \boldsymbol{y})$. The Bayes' factor for source level propositions is then written as follows:

$$V = \frac{f(\boldsymbol{x},\boldsymbol{y}|H_p,I)}{f(\boldsymbol{x},\boldsymbol{y}|H_d,I)}.$$

Bayes' theorem and the rules of conditional probability apply to probability density functions $f(\cdot)$ as well as to probabilities. The value of the evidence V of the evidence may be rewritten – following the argument presented in the section on discrete data – as

$$V = \frac{f(\boldsymbol{y}|\boldsymbol{x},H_p,I)}{f(\boldsymbol{y}|H_d,I)}.$$

This formulation of the expression for V shows that for the numerator the distribution of the measurements on the recovered items, conditional on the measurements on the control items as well as I, is considered. For the denominator, the distribution of the measurements on the recovered items is considered over the distribution of the whole of the relevant population. The denominator is called the 'marginal distribution' of the measurements on recovered items in the relevant population.
In a Bayesian approach, the characteristic of interest is parameterized, for example by the mean. Denote the parameter by θ. This parameter may vary from source (window) to source (another window).

Consider that the two source level propositions to be compared are H_p: the recovered material is from the same source as the control material, and H_d: the recovered material is from a source other than the control material.
The measurements $\boldsymbol{x}$ are from a distribution with parameter θ_1, say and the measurements $\boldsymbol{y}$ are from a distribution with parameter θ_2, say. If $\boldsymbol{x}$ and $\boldsymbol{y}$ come from the same source, then $\theta_1 = \theta_2$, otherwise $\theta_1 \neq \theta_2$. In practice, the parameter θ is not known and the analysis is done with the marginal probability densities of $\boldsymbol{x}$ and $\boldsymbol{y}$. The above equation for V can be rewritten as:

$$V = \frac{\int f(y|\theta) f(x|\theta) \pi(\theta) \mathrm{d}\theta}{\int f(x|\theta) \pi(\theta) \mathrm{d}\theta f(y|\theta) \pi(\theta) \mathrm{d}\theta}.$$

For those unfamiliar with these kinds of manipulations, Bayes' theorem applied to conditional probability distributions is used to write $f(\theta|\boldsymbol{x})$ as $f(\boldsymbol{x}|\theta)\pi(\theta)/f(x)$. The law of total probability with integration replacing summation is used to write $f(x)$ as $f(\boldsymbol{x}|\theta)\pi(\theta)\mathrm{d}\theta.$ Note that $\pi(\theta)$ represents the prior distribution on the

unknown parameter. Therefore, the Bayes' factor does not depend only upon the sample data. It is the ratio of two weighted likelihoods.
Often, the distributions of $(\boldsymbol{x}|\theta)$and $(\boldsymbol{y}|\theta)$ are assumed to be Normal, with θ representing the mean, varying from source to source, and the variance is assumed to be constant from source to source. Those assumptions can be relaxed and (a) various possibilities can be assumed for the distribution of $(\boldsymbol{x}|\theta)$, $(\boldsymbol{y}|\theta)$, and θ and (b) a three-level hierarchical model (variance assumed not constant) can be considered. Moreover, developments for multivariate data are also possible.

Principles of Evidence Evaluation

Three principles arise from the application of the ideas outlined so far.
First, an evaluation is meaningful only when at least one alternative proposition is considered. So, the distribution of the data has to be considered under (at least) two propositions, typically that of the prosecution and that of the defence.
The second principle is that evaluation is based on consideration of probabilities of the evidence E, given a particular issue, H_p, is assumed true, $\Pr\left(E|H_p\right)$ and, given an alternative issue, H_d, is assumed true, $\Pr(E|H_d)$.
The third principle is that the evaluation and interpretation of the evidence is carried out within a framework of circumstances. It has to be conditioned on the background information I.
The application of those principles guarantees some desiderata in the scientist's attitude in evaluating and offering evidence, such as balance, transparency, robustness, and added value. The degree to which the scientist succeeds in meeting these criteria depends crucially on the chosen inferential framework which may be judged by the criteria of flexibility and logic.

Interpretation
Continuous Data and Continuous Variables (Propositions)

So far, the outline focused on categorical (or continuous) data and discrete propositions, but Bayesian analysis also deals with situations involving continuous variables. A typical instance of this is the situation where a parameter, such as a mean, needs to be estimated. As an example, suppose that a random sample, x $= (x_1, \ldots, x_n)$, is available. For example, this may be the case when a scientist is interested in blood alcohol concentration on the basis of a series of n measurements taken from a given individual arrested by traffic police.
Suppose further that the data follow a Normal distribution with unknown mean, θ, and known variance, σ^2. Suppose also that there is some background information available so that some values of θ seem more likely a priori. Then, assuming a conjugate Normal prior distribution for the parameter of interest, that is the mean θ, having a mean μ and a variance τ^2, it can be shown that the posterior density is still a Normal distribution, $N(\mu_x, \tau_x^2)$, with

$$\text{mean } \mu_x = \frac{\frac{\sigma^2}{n}}{\frac{\sigma^2}{n}+\tau^2}\mu + \frac{\tau^2}{\frac{\sigma^2}{n}+\tau^2}\bar{x}, \text{ and variance } \tau_x^2 = \frac{\frac{\sigma^2}{n}\tau^2}{\frac{\sigma^2}{n}+\tau^2}.$$

The posterior mean is a weighted average of the prior mean μ and the sample mean $\bar{x}$, with weights proportional to the variances corresponding to the prior distribution and the sampling distribution. Comparable lines of reasoning can be invoked to approach situations involving unknown variances, alternative distributions and data distributions.

Pitfalls of Intuition

The Bayesian approach to the interpretation of evidence enables various errors and fallacies to be exposed. The most well-known of these are the prosecutor's and defence attorney's fallacies. As an example, consider a crime where a blood stain is found at the scene and it is established that it has come from the criminal. Only for sake of illustration, consider that the stain has a profile which is present in only 1% of the population. It is also supposed that the size of the relevant population is 200 000. A POI is identified by other means and their blood is found to be of the same profile as that found at the crime scene.
The prosecutor argues that, because the blood profile is present in only 1% of the population, there is only a 1% chance that the POI is not the offender. There is a 99% chance that he is the offender. The defence attorney argues that, because 1% of 200 000 is 2000, the POI is only one person in 2000 who have a corresponding blood profile. There is a probability of 1/2000 that he is the offender. This is then used to argue that the blood group is, therefore, of little probative value and not very helpful in the case.

Consideration of the odds form of Bayes' theorem explains these fallacies. Denote the blood evidence by E and let the two competing propositions be H_p, the POI is the offender, and H_d, an unknown person is the offender. Then the odds form of Bayes' theorem is that (omitting *I* from the notation)

$$\frac{\Pr(H_p\,|E)}{\Pr(H_d\,|E)} = \frac{\Pr(E|H_p)}{\Pr(E|H_d)} \times \frac{\Pr(H_p\,)}{\Pr(H_d\,)}.$$

The Bayes' factor is $\Pr(E|H_p)/\Pr(E|H_d) = 1/0.01 = 100$. The posterior odds are increased by a factor of 100. Consider the prosecutor's statement. It claims that the probability of the POI being the offender, after presentation of the evidence, is 0.99. In formal terms, this corresponds to $\Pr(H_p|E) = 0.99$ and, hence, $\Pr(H_d|E) = 0.01$. The posterior odds are 99, which is approximately 100. V is also 100. Thus, the prior odds are 1 and $\Pr(H_p) = \Pr(H_d) = 0.5$. For the prosecutor's fallacy to be correct the prior belief is that the POI is the offender is just the same as an unknown person being the offender.

The defence argues that the posterior probability $\Pr(H_p|E)$ equals $1/2000$ and, hence, $\Pr(H_d|E)$ equals $1999/2000$. The posterior odds are $1/1999$, which is approximately $1/2000$. Since the posterior odds are bigger by a factor of 100 than the prior odds, the prior odds are $1/200\,000$, or the reciprocal of the population size. The defence is arguing that the prior belief in the POI being the offender is approximately $1/200\,000$. This could be expressed as a belief that the POI is just as likely to be liable as anyone else in the relevant population. The fallacy arises because the defence then argues that the evidence is not relevant. However, before the evidence was led, the POI was one of 200 000 people, after the evidence was led he is only one of 2000 people. Evidence which reduces the size of the pool of potential criminals by a factor of 100 is surely relevant.

Other errors have been identified. The 'ultimate issue error' is another name for the prosecutor's fallacy. It confuses the probability of the evidence *if* a POI is not the offender with the probability *that* the POI is not the offender, given the evidence. The ultimate issue is the issue proposed by the prosecution of which it is asking the court to find in favour. The 'source probability error' is to claim the person of interest is the source of the evidence. This would place the POI at the scene of the crime but would not, in itself, be enough to show that he was the offender. The 'probability (another match) error' consists in equating the rarity of a characteristic with the probability that another person has this characteristic. The 'numerical conversion error' equates the reciprocal of rarity of the corresponding characteristic to the number of people that have to be examined before another person with the same characteristic is found.
More generally, high values for the evidence provide strong support for the prosecution's case. They are not, however, sufficient in themselves to render a guilty verdict. The prior odds have to be considered as well. Very high values for the evidence, when combined with very small values for prior odds, may produce small values for the posterior odds. This may be the case when the suspect has been selected as a result of a database search and when there is little or no other evidence against the POI.

(Bayesian) decision theory

Decision theory specifies how to make choices between *courses of action* (also called *decisions*) when there is uncertainty about some event (propositions) or parameter (quantity), called *states of nature*, and about the *consequences* of these choices. Thus, the basic elements characterizing decision theory are the states of nature denoted by θ (which may be discrete or continuous) in the parameter space Θ, the list of exclusive and exhaustive decisions denoted by letter d in a decision space D, and the consequences c that represent the combination of a decision d taken when the state of nature is θ. Consequences are described, in mathematical terms, as $c(d, \theta)$ and the set of all consequences is denoted C.

Uncertainty about the states of nature is quantified through a probability distribution over the parameter space. This is done through Bayes' theorem as new evidence is acquired.
The desirability of each consequence is quantified. Values for consequences are assigned on a numerical scale by the decision maker through a function called the *utility function*, denoted by $\mathrm{U}(\cdot)$. It associates a utility value $\mathrm{U}(d, \theta)$ to each one of the possible consequences $c(d, \theta)$. The utility values clarify which consequence is more desirable than others, and which consequences are considered equivalent. Finally, by combining the utilities $\mathrm{U}(d, \theta)$ associated to each consequence and the probabilities associated to the states of nature, one can quantify a measure of the desirability of each potential decision in terms of what is known as *expected utility*. The value of the expected utility is the criteria for the rational decision maker's choice between courses of action. The decision with the largest expected utility value represents one's best choice, because it is the decision with the highest probability of obtaining the most favorable consequence.

The expected utility value is computed as follows in situation involving discrete states of nature θ:

$$\mathrm{EU}(d) = \sum_{\theta\in\Theta} \mathrm{U}(d,\theta)\Pr(\theta|E,I).$$

Under situations involving continuous states of nature θ, the probability mass function is replaced by a probability density function on the parameter space. The expected utility is given by:

$$\mathrm{EU}(d) = \int_{\theta\in\Theta} \mathrm{U}(d,\theta)f(\theta|E,I)\mathrm{d}\theta.$$

Specialised statistical, legal and forensic literature introduced formal decision theoretic developments for a variety of decisions arising in legal contexts, such as making decisions about ultimate issues, forensic identification (individualization), sampling and classification, among several others.

Conclusion

This chapter emphasises three fundamental aspects. The first is the consideration that probability is the reference framework for scientific reasoning under uncertainty. Second, Bayes' theorem – a result following from the definition of probability and related concepts – provides a way to formalise the assessment of the value of scientific evidence with respect to propositions of judicial interest. Finally, as an extension of probability theory, decision theory provides a framework that can shed light on how to use scientific evidence in coherent decision analysis under uncertainty.

See also

BIOLOGY/DNA | Bayesian Networks; BIOLOGY/DNA | DNA – Statistical Probability;
FOUNDATIONS | Overview and Meaning of Identification/Individualization.

Glossary

Bayes' theorem: Bayes' theorem is a consequence of the basic laws of probability and can be applied for revising beliefs about uncertain propositions in the light of new evidence. In legal contexts, reasoning according to Bayes' theorem is used in order to examine whether particular evidence strengthens or weakens a case. More generally, Bayes' theorem provides a standard for logically correct reasoning under uncertainty.

Likelihood ratio: A likelihood ratio is defined by a ratio of two conditional probabilities: the probability of the evidence given each of two mutually exclusive and competing propositions. In forensic science applications, the likelihood ratio is used as an expression for the meaning of scientific evidence and as a measure for its probative value.

Probability: Probability is a measurement device for uncertainty. In one of its most widespread interpretations, it serves the purpose of expressing an individual's personal degrees of beliefs about uncertain propositions. The laws of probability can be derived in serval different ways, not merely axiomatic considerations, thus constituting a fundamental framework for inductive logic.

Decision theory: A normative theory of decision making that allows decision makers to assess the consequences of alternative courses of action, compare them and ensure coherent (decision) behaviour.

Further Reading

Aitken CGG and Lucy D (2004) Evaluation of trace evidence in the form of multivariate data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 53: 109-122.

Aitken CGG, Taroni F and Bozza S (2021) *Statistics and the Evaluation of Evidence for Forensic Scientists*, 3rd edition, Chichester: John Wiley & Sons.

Biedermann A, Bozza S, and Taroni F (2016) The decisionalization of individualization. Forensic Science International 266: 29-38.

Bozza S, Taroni F, Raymond R, and Schmittbuhl M (2008) Probabilistic evaluation of handwriting evidence: Likelihood ratio for authorship. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 57: 329-341.

de Finetti B, La probabilità : guardarsi dalle contraffazioni. *Scientia* 111 (1976) 225-281.

ENFSI (2015) ENFSI Guideline for evaluative reporting in forensic science, Dublin. http://enfsi.eu/documents/forensic-guidelines/ (last accessed 8 November, 2021).

Evett IW (1984) A quantitative theory for interpreting transfer evidence in criminal cases. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 33: 25-32.

Evett IW (1987) Bayesian inference and forensic science: Problems and perspectives. *The Statistician* 36: 99-105.

Evett IW, Lambert JA, and Buckleton JS (1998) A Bayesian approach to interpreting footwear marks in forensic casework. *Science & Justice* 38: 241-247.

Evett IW and Weir BS (1998) *Interpreting DNA Evidence*. Sunderland, MA: Sinauer.

Good IJ (1991) Weight of evidence and the Bayesian likelihood ratio. In: Aitken CGG and Stoney DA (eds.) *The Use of Statistics in Forensic Science*, pp. 85-106. Chichester: John Wiley & Sons.

Kaye DH (1988) What is Bayesianism? In: Tillers P and Green ED (eds.) *Probability and Inference in the Law of Evidence - The Uses and Limits of Bayesianism*, pp 1-19. Dordrecht: Springer.

Koehler JJ, Chia A, and Lindsey S (1995) The random match probability in DNA evidence: Irrelevant and prejudicial? *Jurimetrics Journal* 35: 201-219.

Lindley DV (1977) A problem in forensic science. *Biometrika* 64: 207-213.

Lindley DV (1985) *Making Decisions*, 2nd edition. Chichester: John Wiley & Sons.

Robertson B, Vignaux GA and Berger CEH (2016) *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*, 2nd edition. Chichester: John Wiley & Sons.

Savage LJ (1954) *The Foundations of Statistics*. New York : Wiley.

Schum DA (2001) *Evidential Foundations of Probabilistic Reasoning*. Evanston: Northwestern University Press.

Taroni F, Bozza S, Biedermann A, Garbolino P, and Aitken CGG (2010) *Data Analysis in Forensic Science: A Bayesian Decision Perspective*. Chichester: John Wiley & Sons.

Taroni F, Biedermann A, Bozza S, Garbolino P, and Aitken CGG (2014) *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science*, 2nd edition. Chichester: John Wiley & Sons.

Taroni F, Bozza S, Biedermann A (2021), Decision theory, in: *Handbook of Forensic Statistics*, Banks DL, Kafadar K, Kaye DH, Tackett M (Eds.), Boca Raton: Chapman & Hall/CRC:103-130.

Thompson WC and Schumann EL (1987) Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defence attorney's fallacy. *Law and Human Behaviour* 11: 167–187.