Evidence and conviction:

Rational reasoning since Aristotle

Charles Berger and Diederik Aben

Dr. ir. C.E.H. Berger is principal scientist at the Netherlands Forensic Institute, and member of the Board of the Netherlands Register for Court Experts (NRGD).

Mr. D.J.C. Aben is Advocate General at the Supreme Court of the Netherlands, and editor of this journal.

Judges sometimes say they do their work based on Aristotle, experience and intuition. Gut feeling, so to speak. Science would thus have little to do with this, and have little to offer to the judge that reaches his verdicts by using his experience. Science however, has evolved since the times of Aristotle (384-322 BC). This is no different for forensic science. Can the judge do without?

This article is the first of a series of three, in which the authors dwell on the possibilities of using a Bayesian reasoning scheme in evaluating evidence, and common fallacies. In this part the historical development of the reasoning scheme.

Introduction

The criminal trial is based on the premise that the truth not only exists, but can actually be found. In any case it must be sought in the trial, and then a final decision should be taken. This means that in the judicial investigation relevant data are collected, after which the selection and evaluation is a duty of the court. Finally, a verdict follows in which a decision is taken on the scenario that the prosecutor presented to the court in the form of the indictment. Whatever the outcome of the trial will be, all participants should have had the opportunity to discuss the collected information before a decision is taken.

To that extent, there are striking similarities between this judicial process and the practice of science. Scientists also collect data, explore theories (scenarios) and do not eschew debate.

But there are also differences. Scientists have their reservations about the somewhat naive word ‘truth’, and they frown even more at something like ‘the whole truth.’ Scientific research is based on the collection of empirical data, and formulating and testing hypotheses on the basis of such data. The scientist does not have the task to take a decision. Even if he chooses for one particular hypothesis, that choice will be provisional and new observations may still require him to reject that hypothesis.

Judges however, must at some point reach a decision in the form of a verdict that is expressed in absolute terms. The crime is proven or not. The judge can not pronounce in his verdict that the accused is *likely* or *unlikely* to have committed the crime. Judges are also not in a position to reconsider earlier choices. The usual procedures of an appeal and further appeal are the only way to accomplish that. Moreover, the trial must be completed within a reasonable time. That way, it is difficult to make use of years of advancing insight [1].

Lawyers are not used to the concept of a ‘hypothesis’ within the scope of the truth-finding process. When Prof. Dr. R. Meester, professor of probability and statistics associated with the VU university of Amsterdam, had dropped that term more than once during his questioning at a hearing as an expert, he was corrected by the judge: ”Mr. Meester, we do not engage in hypothesizing, but in truth finding!”.

A key point is that in fact judges *do* make use of hypotheses and probabilities in their quest for the truth. And it is a good thing that they do so. The hypotheses include the scenarios regarding the offense that is the subject of investigation, and are discussed by the prosecution and the defense. Or – of considerable importance – the scenarios that the court may consider themselves. Reasoning with all hypotheses and conditions established, the court ultimately reaches its decision; at least that is the model that we envisage and of which we suspect that the judge uses it, whether intentionally or not.

The judge cannot avoid reasoning with information which is inherently uncertain. Until he decides whether the total of the data is sufficient for a conviction, he will not escape probabilistic reasoning in that sense. We believe that insights from science can be extremely helpful to the court. From several perspectives, we will discuss how a Bayesian reasoning scheme can help truth-finding in the judicial process. It is closely related to how forensic scientists should do their work, and to how they phrase their conclusions. If only for that reason, the criminal courts should have some knowledge of this scheme. But the reasoning scheme is just as important to the reasoning of the other participants in the legal process, and – contrary to what is sometimes thought – its use does not depend on the availability of scientific data and hard numbers.

History

Aristotle wrote his Rhetoric [2], a masterpiece on the art of persuasion, between 360 and 330 BC. He made a distinction between political, ceremonial and forensic rhetoric. Forensic rhetoric was aimed at convincing a judge that had to render a judgment on past events. Aristotle realized that the propositional and predicate logic he had worked out, which focused on arguments with hard facts, would rarely be useful for that:

> *'There are few facts of the 'necessary' type that can form the basis of rhetorical syllogisms. Most of the things about which we make decisions, and into which therefore we inquire, present us with alternative possibilities. For it is about our actions that we deliberate and inquire, and all our actions have a contingent character; hardly any of them are determined by necessity. Again, conclusions that state what is merely usual or possible must be drawn from premises that do the same, just as 'necessary' conclusions must be drawn from 'necessary' premises'.*

So Aristotle realized – even then – that in addition to the logic that with proper premises yields necessarily correct conclusions, arguments exists with premises that are likely but not necessarily correct. Despite the lack of necessary correctness, these arguments should influence the conviction of a rational person.

> *'The argument may, for instance, be that Dionysius, in asking as he does for a bodyguard, is scheming to make himself a despot. "For in the past Peisistratus kept asking for a bodyguard in order to carry out such a scheme, and did make himself a despot as soon as he got it; and so did Theagenes at Megara"; and in the same way all other instances known to the speaker are made into examples, in order to show what is not yet known, that Dionysius has the same purpose in making the same request: all these being instances of the one general principle, that a man who asks for a bodyguard is scheming to make himself a despot.*
>
> *We have now described the sources of those means of persuasion which are popularly supposed to be demonstrative.'*

Here, examples from the past – known to Aristotle – are used to make an assertion about an unknown, uncertain future. That assertion is not necessarily correct but the examples can influence the conviction of a rational person. In itself this is not strange observation. In everyday life, people gather information derived from observations of examples all the time. At the time however, this was a groundbreaking way of thinking, which was forgotten for a long time after Aristotle.

In order to expand his logic from certain to uncertain events, Aristotle needed probability theory. But the development of the concept of probability would take another two millennia. Until then, logical reasoning had to be limited to either black and white, dichotomous reasoning, or non-scientific reasoning.

For deductive reasoning as with Aristotle's "syllogism" there was no problem: All men are mortal and Socrates is a man, therefore Socrates is mortal. In such an argument the conclusion follows necessarily from the previous premises. The logic known at the time was sufficient to provide for the formalization of this argument. But to come to new knowledge more was needed, a different type of reasoning: inductive reasoning.

John Stuart Mill wrote about it in the 19th century (1859) in *A system of logic, ratiocinative and inductive* [3]:

> *'In every induction we proceed from truths which we knew, to truths which we did not know: from facts certified by observation, to facts which we have not observed, and even to facts not capable of being now observed; future facts, for example: but which we do not hesitate to believe upon the sole evidence of the induction itself.'*

Induction is exactly the kind of reasoning with which Aristotle's example of the guard can be understood. It does not matter whether we are "predicting" the unknown past or the future: it is the evidence of our observations regarding competing hypotheses about what could not be observed directly that matters.

Mill has described how the most likely cause of an observed effect (result) can be found when the hypotheses about the cause are equally probable before observing the effect:

> *'Common sense and science alike dictate that, all other things being the same, we should rather attribute the effect to a cause which if real would be very likely to produce it, than to a cause which would be very unlikely to produce it.'*

With that, he went much farther than Aristotle, who spoke only on arguments about cause and effect that could convince a rational person. Aristotle lacked a sufficient concept of probability to be able to clarify why such a argument has a certain persuasive power. Mill however, examined from what this persuasive power arose. He concluded that the conviction increases if the result is more likely when one hypothesis is true, than when the other hypothesis is true.

Meanwhile, probability theory had been introduced, initially mainly for the analysis of betting and card games (see *e.g.* Christiaan Huygens [4] (1629-1695) in

1657). Scientists such as Pierre Simon de Laplace (1749-1827) applied probability theory (or "Doctrine of Chances") to scientific and practical problems. In the introduction to his *Théorie analytique des probabilités* [5] (1814) he tells us how far-reaching the importance of his work is:

> *'In this introduction I will present the principles of probability theory, and the results to which I have come in this work by applying them to the key questions of life, which in fact, are mostly probability problems. You can even say that, strictly speaking, almost all our knowledge is only probable'*

In this work he indicates somewhat cryptically how from the observation of an effect, the probability of possible causes can be calculated:

> *Any cause to which an observation can be attributed, is indicated with as much more probability as the probability of that observation when the cause is supposed to exist;*

Noah K. Davis (1830-1910) describes the same a lot clearer in his Elements of inductive logic (1895) [6]:

> *'Given an effect to be accounted for, and there being several causes that might have produced it, but of whose presence in the particular case nothing is known; the probability that the effect was produced by any of these causes is as the antecedent probability of the cause, multiplied by the probability that the cause, if it existed, would have produced the given effect.'*

The observation of an effect is thus evidence (*E*) for a hypothesis about the cause (*H*). In mathematical notation, Laplace and Davis say that in this situation:

$$P(H|E) = P(H) \cdot P(E|H).$$

In this notation $P(H|E)$ means: the probability of hypothesis H being true, given the observation of the evidence E. And $P(E|H)$ means the probability of observing E, given that hypothesis H is true.[7]

Poisson (1781-1840) writes in 1837 [8] on Laplace's work:

> *'The solution he gave of this problem, one of the most delicate of probability theory, is based on the principle that serves to determine the probabilities of diverse causes to which one can attribute the observed facts; a principle that Blayes [sic] has presented first in a slightly different form, and of which Laplace subsequently has made the best use, in his memoirs and his treatise, for calculating the probability of future events after the observation of past events'.*

Poisson thus refers to the theorem named after an English clergyman, Thomas Bayes (1702-1761), whose work on this subject was published posthumously in 1763 [9] and contained a special case of this theorem. The work commenced by Bayes was continued by Laplace.

A very helpful form of Bayes' theorem is expressed in odds: the probability ratios that we know from gambling on *e.g.* the outcome of horse races. This formula shows how an observation of an effect provides evidence E to help us choose between two possible causes (hypotheses H_1 and H_2):

$$\underbrace{\frac{P(H_1)}{P(H_2)}}_{\text{prior odds}} \times \underbrace{\frac{P(E|H_1)}{P(E|H_2)}}_{\text{likelihood ratio}} = \underbrace{\frac{P(H_1|E)}{P(H_2|E)}}_{\text{posterior odds}}$$

The prior and posterior odds are a good measure for our degree of belief in the truth of either hypothesis before (prior) and after (posterior) observation. The likelihood ratio (LR) is a measure of the increase in this conviction, and therefore of the evidential value of the observation [10]. The court is ultimately interested in the posterior odds: the probabilities of the hypotheses given the observation.

In the formula, we see how the odds are increased by the observation (E): by a factor (the LR) equal to the probability of that observation when Hypothesis 1 is true, divided by the probability of that same observation when Hypothesis 2 is true. This brings us one step further than the qualitative indication of J.S. Mill.

As an example, let us apply this knowledge to the case mentioned by Aristotle of the despot who asks for a bodyguard.

In it, the evidence is the observation that Dionysius asks for a bodyguard. What does this observation tell us about the hypotheses that are to be tested? There are two hypotheses:

H_1 = Dionysius will become a despot;

H_2 = Dionysius will not become a despot.

On the basis of other, previously established data, we could estimate the prior odds: our degree of belief prior to the observation of the bodyguard request.

Bayes' theorem tells us that due to the observation the prior odds increase by

$$\frac{P(E \mid H_1)}{P(E \mid H_2)}.$$

To estimate the evidential value of the observation, we must not only examine the frequency with which despots asked for the formation of a private bodyguard prior to their despotism, but also how often non-despots did the same. Aristotle explicitly mentions two examples of the first: Peisistratus and Theagenes. For a better estimate, we would obviously like to know how a larger number of non-despots and despots have acted prior to their regimes, but we'll have to do with the data that Aristotle furnishes.

(1) Suppose that prior to the bodyguard request we estimate the probability that Dionysius will be a despot at 75%, based on other information. The prior odds then are

$$\frac{3/4}{1/4} = 3$$

(3 against 1).

(2) Suppose that in the past there were three despots, two of which (Peisistratus and Theagenes) previously asked for a bodyguard (2 out of 3). Of the six non-despots from the past, two have also asked for bodyguards (2 out of 6), without subsequently developing into a despot. From those ancient observations in this example, we will derive the strength of the evidence.

Our prior odds (of 3 to 1) increase due to those observations by a factor equal to

$$\frac{2/3}{2/6} = 2.$$

This factor of 2 is the likelihood ratio (LR) in this example. In a medical context the synonymous term ‘diagnostic value’ is often used. In this case this ratio indicates the evidential value of the observation that Dionysius asks for a bodyguard. The worrying request for a bodyguard makes the odds for despotism double from 3 to 6. Those are the posterior odds, in which the evidence is taken into account. They are 6 against 1 that Dionysius will be despot. The estimation of the probability that he will become a despot will therefore increase from 75% to 86% (6/7). With this, the millennia old example of Aristotle was solved! [11]

Time to take stock.

In the first place, it’s important to realize that the evidential value of the observation is related to at least two hypotheses. In the second place, note that the evidential value (likelihood ratio) is relative: only the increase (or decrease) of our degree of belief is given, and not the degree of belief itself. The degree of belief given the evidence is expressed by the posterior odds, for which in addition to the evidence, the prior odds need to be taken into account. Let’s reflect on this for a moment.

It is a common mistake to think that the probability of a hypothesis can be directly derived from (the evidential value of) an observation. A historical example is the miscarriage of justice which has become known as the Dreyfus affair. Among the evidence against the French Jewish army officer Dreyfus (1859-1935) were documents that were attributed to him – wrongly, as it turned out – and from which would follow that he was active as a spy for the Germans. The prominent forensic expert Alphonse Bertillon (1853-1914) committed this mistake in the interpretation of the handwriting evidence, a point that was made painfully clear in 1906, by the renowned mathematician Henri Poincaré (1854-1912) [12]:

> *'... given the impossibility of knowing the prior probability, we cannot say: this agreement proves that the ratio of the probability of a forgery to the inverse probability has this or that value. We can only say, by observing the agreement: that ratio becomes this much larger than before the observation.'*

Without knowing the prior odds we can't say anything about the posterior odds. If we know the likelihood ratio (the evidential value), we only know by how much our degree belief should increase or decrease but not what that degree of belief should be.

It is also possible that the evidence is not informative at all on the probability of the hypotheses being true. Indeed when the observation is equally likely under both hypotheses, the evidential value is equal to 1. The observation does not increase or decrease our degree of belief, and in this case is irrelevant for the hypotheses considered.

Let's go back to our 'classic' example.

If in Aristotle's example Peisistratus had not become despot, then the LR would have been equal to

$$\frac{1/3}{2/6} = 1.$$

An observation with an evidential value of 1 does not distinguish between the hypotheses and is neutral, or in other words, of no weight.

Such a criterion for relevancy can be recognized in the U.S. Federal Rules of Evidence:

> *'Rule 401. Definition of "Relevant Evidence"*
>
> *"Relevant evidence" means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.'*

In the above we were making calculations with numbers. But when no quantitative data are available, the same reasoning is equally applicable.

The core of the above can be summarized as follows:

1. The evidential value of observations ('evidence') is related to the hypotheses tested. This concerns at least two hypotheses whose probabilities are considered relative to one another. The Bayesian reasoning scheme thus forces us to take alternatives into consideration, and – in a criminal justice context – not merely the account of events of *e.g.* the prosecution.

2. An observation ('evidence') points in the direction of the hypothesis under which the observation is most likely. That "pointing" is relative: the ratio of the probabilities of the observation under both hypotheses (the LR) gives the relative increase of our degree of belief, regardless of how large or small that belief was before the observation.

3. The evidential value in itself says nothing about the probability of the hypotheses without knowing, or without at least making assumptions about the probability of those hypotheses prior to knowing the evidence, *i.e.*: the prior odds. In one line: no posterior odds without prior odds.

There is one more point to make. Not all knowledge is uncertain and not all statements are probability statements. Sometimes something is obviously not true. There are indeed factual and logical impossibilities which we may safely assume did not occur. We only give a reassuring example: none of the authors of this article has given Socrates the hemlock.

In the next part of this triptych, entitled “Reasoning in the courtroom”, we will discuss the application of the Bayesian reasoning scheme in criminal justice. Therein we will demonstrate the above using examples. It will not require calculations. The core of our argument is that the Bayesian reasoning scheme promotes insight, even without quantifying the evidence and the odds.

The triptych “Evidence and conviction” is about rational reasoning for scientists and lawyers. The Bayesian framework gives a number of rules and provides a clear view of the pitfalls.

1. Evidence and conviction: Rational reasoning since Aristotle

2. Evidence and conviction: Reasoning in the courtroom

3. Evidence and conviction: A clear view of the pitfalls

References

[1.] Even the extraordinary remedy of a revision offers little or no solace, because changed scientific insights are difficult to translate into a novelty.

[2.] Marc Huys, *Aristoteles Retorica*, Historische Uitgeverij 2004.

[3.] John Stuart Mill, *A system of logic, ratiocinative and inductive, being a connect ed view of the principles of evidence and the methods of scientific investigation*, Harper & Brothers Publishers 1859

[4.] Christiaan Huygens' *De Ratiociniis in Ludo Aleæ* was published in Latin in 1657.

[5.] Laplace, *Théorie analytique des probabilités,* Courcier 1814.

[6.] Noah K. Davis, *Elements of Inductive Logic*, Harper & Brothers Publishers 1895.

[7.] These two probability statements should not be confused. Think for example of the probability that a woman is pregnant. That chance is not equal to the probability that a pregnant person is... a woman.

[8.] Siméon-Denis Poisson, *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile*, Bachelier 1837.

[9.] Thomas Bayes, 'An Essay towards solving a Problem in the Doctrine of Chances', *Philosophical Transactions of the Royal Society of London* 53 (1763), 370-418.

[10.] This observation can be a measurement result, but also a judicial observation, or an observation by a witness.

[11.] Dionysius (c. 430-367 BC) in reality *did* become a despot.

[12.] Henry Mornard, *L'affaire Dreyfus: la revision du procès de Rennes*, Ligue Française pour la défense des droits de l'homme et du citoyen 1907, p. 334.