Improving statistical estimates used in the courtroom

Professor Norman Fenton

Queen Mary University of London and Agena Ltd

Address:
Queen Mary University of London
School of Electronic Engineering and Computer Science
London E1 4NS.
Email: norman@eecs.qmul.ac.uk

Tel: 07932 030084

Precis

Misunderstanding and misuse of statistics in the law is a continuing concern. The many dozens of widely documented cases [3][7] in which verdicts have been influenced by incorrect statistical reasoning represent only the tip of the iceberg.

Most common fallacies of statistical reasoning are easily avoided by using the Bayesian approach, which provides the only rational way to evaluate the weight of evidence. Yet Bayes is especially widely misunderstood and mistrusted. A recent Court of Appeal Ruling (known as RvT) [1] dealt a blow to this approach by appearing to rule it out as an admissible way to present expert evidence in all but a narrow, ill-defined, class of forensics. The effect is that relevant evidence from experts may be suppressed, and there may be more, not less, confusion due to poor statistical reasoning. To confront this challenge we have formed an international consortium of influential forensic scientists, statisticians and academic and practising lawyers to determine explicit guidelines for when and how Bayesian reasoning should be used in presenting evidence.

Bayes' Theorem

Bayes' theorem is a basic rule for updating the probability of a hypothesis given evidence. Most importantly, from a legal perspective, Bayes' theorem tells us how to weight the impact of pieces of evidence individually or in combination. Suppose, for example, that blood (the type of which is prevalent in one in every thousand people) is found at the scene of a crime and the defendant has the same blood type. Clearly this piece of evidence is important because it increases our belief in the probability that it really was the defendant's blood at the scene. But by how much? To answer this we need to compare the prosecution 'likelihood' (the probability of seeing the evidence if the prosecution hypothesis is true), with the defence likelihood (the probability of seeing the evidence if the defence hypothesis is true). The prosecution likelihood can be assumed to be equal to one, while the defence likelihood is one in a thousand. So we are 1000 times more likely to observe the evidence if the prosecution hypothesis is true than if the defence hypothesis is true.

The ***likelihood ratio,*** the prosecution likelihood divided by the defence likelihood (1000 in this case), is therefore a simple measure of the impact of the evidence. Values above one favour the prosecution hypothesis (the higher the better), while values below one favour the defence hypothesis (the lower the better). A value of one means the evidence is worthless , since it does nothing to shift the balance of probabilities. The likelihood ratio is extremely valuable, but does not enable us to draw any definitive conclusions about which hypothesis is most likely. To do that we need Bayes' theorem. This tells us that, whatever our prior odds were for the prosecution hypothesis, the result of seeing the evidence is such that those odds are multiplied by the likelihood ratio:

$$\text{Posterior odds} = \text{Likelihood ratio} \times \text{Prior odds}$$

Suppose, for example, there were 10,000 other adults living in the area who, like the defendant, could have been at the scene of the crime, i.e. our prior odds are 10,000 to 1 against the prosecution hypothesis. According to Bayes' the correct revised belief once we see the evidence is that the odds still favour the defence, but they have come down to 10 to 1 against the prosecution:

	Prior odds		Likelihood ratio		Posterior odds
Prosecutor	1	×	1000	=	1
Defence	10000		1		10

Equivalently the probability the defendant was not at the scene has gone from 99.99% to about 91%.

Note that the (revised) probability of the defence hypothesis (91%) is very different from the defence likelihood (one in a thousand). A common error is to assume that the two probabilities are equal, such as if a prosecutor states that 'the probability the defendant was not at the scene given this evidence is one in a thousand' . This is the ***prosecutor fallacy*** - one of the most common of the statistical legal fallacies. In this case there is a simple intuitive 'proof' of Bayes (see Figure 1): we simply note that out of the 10,000 other adults we would expect about 10 to have the same blood type as the defendant. The blood match tells us that the defendant is therefore now one of 11 who could have been at the scene. One of the major problems for non-mathematicians in understanding Bayes is that there is rarely such a simple intuitive explanation.

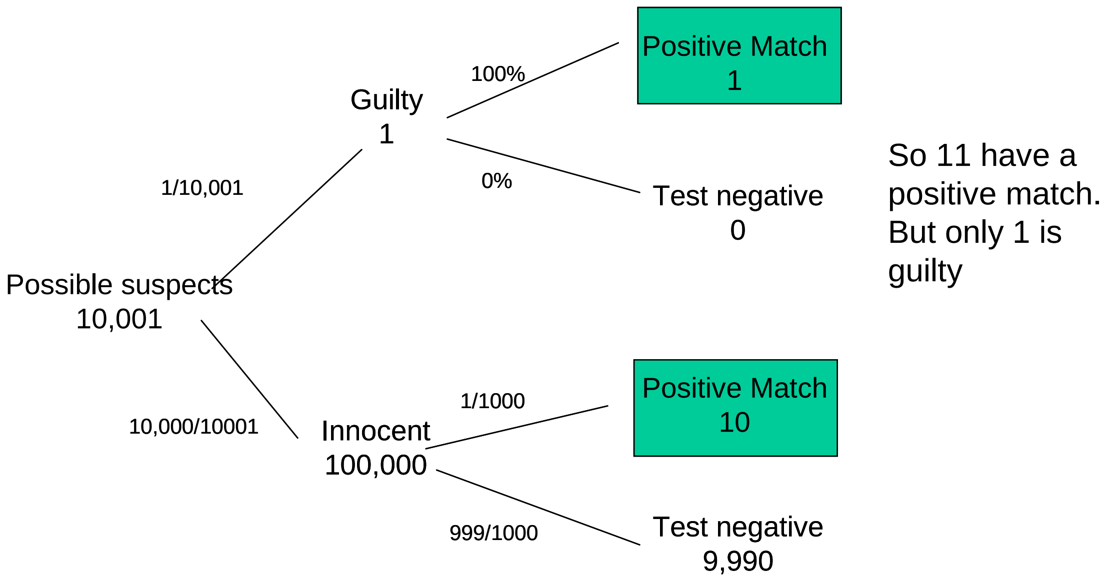

Figure 1 Visual explanation of Bayes. But in practice the problem is much more complex and this diagrammatic approach does not scale up

Resistance to Bayes

Given its potential usefulness in helping to weigh the value of evidence, why is Bayes perceived with such mistrust by much of the legal profession? There are two primary reasons:

1. Misunderstanding of the role of subjectivity

While subjective judgement about uncertain hypotheses is central to the entire jury trial process, lawyers are fiercely resistant to the idea that anybody should provide a numerical figure to describe their subjective uncertainty. But, if we want to use Bayes to turn a likelihood ratio for some evidence into an indication of innocence or guilt then we are indeed forced to provide a prior probability, which may be subjective.

In practice concerns about assigning such priors can be allayed by considering a wide range of possible values. Suppose, for example, that the defendant is known to have been part of a mob of people, exactly one of whom committed the attack. Suppose that a reliable eyewitness says there were about 50 people other than the defendant in the mob. Then a reasonable subjective prior for H ("defendant committed attack") would be 50 to 1 against. It is clearly subjective - there is no 'scientific' data on this. Bayesians would always seek to include a subjective prior such as this, whereas most lawyers would instinctively disallow it. A rational response to the lawyers would be to focus on the potential for disagreement. If the defence finds another witness who claims there were as many as 100 other people in the mob, then it is rational to consider the range from 50 to 1 to 100 to 1. If the likelihood ratio for some piece of evidence (as in the above example) was 1000 then in either case the posterior odds strongly favour the prosecution hypothesis (with the range being 20 to 1 on to 10 to 1 on):

	Prior odds		Likelihood ratio		Posterior odds
Prosecutor	1	×	1000	=	20
Defence	50		1		1
Prosecutor	1	×	1000	=	10
Defence	100		1		1

We would not be able to make this important conclusion if we did not allow subjective priors.

This was the strategy successfully used in a medical negligence case [6] where we evaluated the evidence with Bayes theorem using first the claimant's most extreme assumptions as priors and secondly using the defendant's most extreme assumptions. Even using these extremes Bayes showed the evidence strongly favoured the claimant's hypothesis, hence we could definitively state that the claimant hypothesis was most likely. It follows that, although Bayes' theorem does require prior (possibly subjective) probabilities, it may be sufficient to consider a range of such probabilities that are acceptable to both sides. If the experts don't do this explicitly, then it's highly likely individual jury members will use their own prior assumptions as the basis of their decision implicitly and draw mathematically illogical conclusions.

Misunderstandings of subjectivity are also critical in the controversy surrounding the narrow limits within which it is assumed that Bayes and likelihood ratios can be applied [8]. The judge in RvT ruled they should not be used in evaluating forensic evidence, except for DNA and 'possibly other areas where there is a firm statistical base'. But, ultimately, all probabilities – including even DNA match probabilities from the most comprehensive database – involve some subjective judgement [10]. It is better to embrace subjectivity as inevitable and deal with the differences in subjective judgements explicitly through Bayes.

2. Technical difficulties in presenting Bayesian arguments in full

Only extremely simple Bayesian arguments can be explained from first principles in a way that is understandable to lay people [8]. In any real case there will be more complex assumptions than those in our simple examples; there may even be multiple pieces of dependent evidence. In such cases, the intuitive approaches (such as using population diagrams or decision trees) simply do not scale up. The necessary Bayesian calculations involve formulas that often cannot be computed manually even by the best statistical experts. Instead, they use software tools [7]. Yet, in the case of R v Adams [2] the defence probability expert presented the Bayesian calculations to the jury from first principles. The exercise was, not surprisingly, less than successful and the appeal judge ruled against such use of Bayes in future trials [7].

Getting a scientific consensus

The RvT ruling has drawn fierce criticism from many experts, e.g. [5], who regard it as a constraint on accepted scientific practice and a serious blow to the beneficial use of Bayes in the law. Yet the ruling also confirms widely held views among lawyers. Moreover, even critics of the ruling have praised the attempt to rule out probabilistic forensic evidence based on 'unscientific' data. The fact that there is no definitive notion of what is 'scientific' leads to lawyers erring on the side of caution. For example, as a direct result of R v T, lawyers are now rejecting the use of likelihood ratios even in areas of forensic match evidence where such use was previously standard practice. This means that experts who might previously have shown that the likelihood ratio of some match evidence was between, say, 1000 and 5000, are instead only able to make vague assertions like ' the evidence provides support for the prosecution hypothesis'. Such restrictions represent a backward step for justice.

Although there have been isolated activities such as [4] aimed at improving understanding of probability within the law, it is unlikely that these alone will effect any significant changes on legal practice. Such change can only come about when a critical mass of experts, strongly supported by key members of the judiciary in different countries, are able to reach a consensus about:

1. When Bayesian reasoning about evidence can and cannot be practically applied; and
2. How to get the results of Bayesian analysis accepted in courts without having to present the calculations from first principles.

Wider acceptance of Bayesian analysis also requires lawyers and expert witnesses to understand that there is a crucial difference between:

a) the genuinely disputable (subjective) prior assumptions ; and

b) the Bayesian calculations required to compute the conclusions based on the different disputed assumptions.

By considering ranges of subjective assumptions in a) we can address and overcome the most persistent traditional objections to the use of Bayes. Ultimately this means getting the lay observers to 'accept' that they need only question the prior assumptions that go into the Bayesian calculations and not the calculations themselves. The increasing availability of Bayesian calculation tools (including Bayesian network tools that enable us to model complex chains of interrelated evidence [7]) means that the calculations in b) should never need to be explained in court once statisticians convince lawyers that the results are based on standard proven algorithms .

Conclusion

Proper use of probabilistic reasoning has the potential to improve dramatically the efficiency, transparency and fairness of the criminal justice system. In particular, Bayes' theorem determines how to update the probability of prior beliefs given a piece of evidence. Probabilities are either combined using this rule, or they are combined wrongly. Bayesian reasoning can help experts formulate accurate and informative opinions; help courts determine admissibility of evidence; help identify which cases should and should not be pursued; and help lawyers to explain, and jurors to evaluate, the weight of evidence during a trial. It can also help identify errors and unjustified assumptions entailed in expert opinions.

Unfortunately, there is still widespread disagreement about the kind of evidence to which Bayesian reasoning should be applied and the manner in which it should be presented. While we have suggested methods to overcome these technical barriers, there are still massive cultural barriers between the fields of science and law that will only be broken down by achieving a critical mass of relevant experts and stakeholders, united in their objectives. We are building toward this critical mass of experts [9]. The work must be undertaken by an interdisciplinary team and with the genuine commitment and goodwill of all the relevant professional communities . If we can meet these challenges then there is no reason why Bayes should not take its rightful place as a standard method for evaluating evidence in legal reasoning.

Acknowledgements

This work has benefited from major contributions by Amber Marks, Martin Neil, and Rosie Wild.

References

[1] R v T (2010) EWCA Crim 2439 http://www.bailii.org/ew/cases/EWCA/Crim/2010/2439.pdf

[2] R v Adams [1996] 2 Cr App R 467, [1996] Crim LR 898, CA and R v Adams [1998] 1 Cr App R 377

[3] Aitken, C.G.G. and Taroni, F., Statistics and the evaluation of evidence (2nd Edition). 2004 John Wiley & Sons, Ltd.

[4] Aitken, C., P. Roberts and G. Jackson (2010). Fundamentals of Probability and Statistical Evidence in Criminal Proceedings, Practitioner Guide No 1, Royal Statistical Society's Working Group on Statistics and the Law.

[5] Berger, C. E. H., J. Buckleton, C. Champod, I. Evett and G. Jackson (2011). "Evidence evaluation: A response to the court of appeal judgement in R v T." Science and Justice 51: 43-49

[6] Fenton, N. and M. Neil (2010). "Comparing risks of alternative medical diagnosis using Bayesian arguments." Journal of Biomedical Informatics 43: 485-495.

[7] Fenton, N.E. and Neil, M (2011)., 'Avoiding Legal Fallacies in Practice Using Bayesian Networks', to appear Australian Journal of Legal Philosophy. Draft available here: www.eecs.qmul.ac.uk/~norman/papers/fenton_neil_prob_fallacies_June2011web.pdf

[8] Fenton, N.E. and Neil, M., 'On limiting the use of Bayes in presenting forensic evidence', submitted Forensic Science International 2011. Extended draft here: www.eecs.qmul.ac.uk/~norman/papers/likelihood_ratio.pdf

[9] https://sites.google.com/site/bayeslegal/

[10] Thompson, W. C., F. Taroni and C. G. G. Aitken (2003). "How the probability of a false positive affects the value of DNA evidence." Journal of Forensic Sciences 48(1): 47-54.