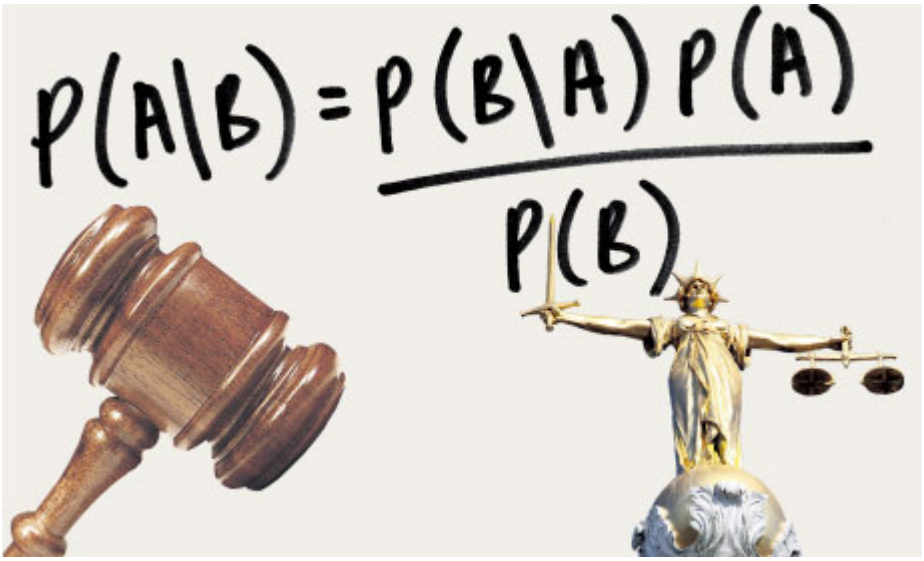
A formula for justice

Bayes' theorem is a mathematical equation used in court cases to analyse statistical evidence. But a judge has ruled it can no longer be used. Will it result in more miscarriages of justice?

Angela Saini
The Guardian, Sunday 2 October 2011 16.30 EDT



Bayes' theorem. Photograph: guardian.co.uk

It's not often that the quiet world of mathematics is rocked by a murder case. But last summer saw a trial that sent academics into a tailspin, and has since swollen into a fevered clash between science and the law.

At its heart, this is a story about chance. And it begins with a convicted killer, "T", who took his case to the court of appeal in 2010. Among the evidence against him was a shoeprint from a pair of Nike trainers, which seemed to match a pair found at his home. While appeals often unmask shaky evidence, this was different. This time, a

mathematical formula was thrown out of court. The footwear expert made what the judge believed were poor calculations about the likelihood of the match, compounded by a bad explanation of how he reached his opinion. The conviction was quashed.

But more importantly, as far as mathematicians are concerned, the judge also ruled against using similar statistical analysis in the courts in future. It's not the first time that judges have shown hostility to using formulae. But the real worry, say forensic experts, is that the ruling could lead to miscarriages of justice.

"The impact will be quite shattering," says Professor Norman Fenton, a mathematician at Queen Mary, University of London. In the last four years he has been an expert witness in six cases, including the 2007 trial of Levi Bellfield for the murders of Marsha McDonnell and Amelie Delagrange. He claims that the decision in the shoeprint case threatens to damage trials now coming to court because experts like him can no longer use the maths they need.

Specifically, he means a statistical tool called Bayes' theorem. Invented by an 18th-century English mathematician, Thomas Bayes, this calculates the odds of one event happening given the odds of other related events. Some mathematicians refer to it simply as logical thinking, because Bayesian reasoning is something we do naturally. If a husband tells his wife he didn't eat the leftover cake in the fridge, but she spots chocolate on his face, her estimate of his guilt goes up. But when lots of factors are involved, a Bayesian calculation is a more precise way for forensic scientists to measure the shift in guilt or innocence.

In the shoeprint murder case, for example, it meant figuring out the chance that the print at the crime scene came from the same pair of Nike trainers as those found at the suspect's house, given how common those kinds of shoes are, the size of the shoe, how the sole had been worn down and any damage to it. Between 1996 and 2006, for example, Nike distributed 786,000 pairs of trainers. This might suggest a match doesn't mean very much. But if you take into account that there are 1,200 different sole patterns of Nike trainers and around 42 million pairs of sports shoes sold every year, a matching pair becomes more significant.

The data needed to run these kinds of calculations, though, isn't always available. And this is where the expert in this case came under fire. The judge complained that he couldn't say exactly how many of one particular type of Nike trainer there are in the country. National sales figures for sports shoes are just rough estimates.

And so he decided that Bayes' theorem shouldn't again be used unless the underlying statistics are "firm". The decision could affect drug traces and fibre-matching from clothes, as well as footwear evidence, although not DNA.

"We hope the court of appeal will reconsider this ruling," says Colin Aitken, professor of forensic statistics at the University of Edinburgh, and the chairman of the Royal Statistical Society's working group on statistics and the law. It's usual, he explains, for forensic experts to use Bayes' theorem even when data is limited, by making assumptions and then drawing up reasonable estimates of what the numbers might be. Being unable to do this, he says, could risk miscarriages of justice.

"From being quite precise and being able to quantify your uncertainty, you've got to give a completely bland statement as an expert, which says 'maybe' or 'maybe not'. No numbers," explains Fenton.

"It's potentially very damaging," agrees University College London psychologist, Dr David Lagnado. Research has shown that people frequently make mistakes when crunching probabilities in their heads. "We like a good story to explain the evidence and this makes us use statistics inappropriately," he says. When Sally Clark was convicted in 1999 of smothering her two children, jurors and judges bought into the claim that the odds of siblings dying by cot death was too unlikely for her to be innocent. In fact, it was statistically more rare for a mother to kill both her children. Clark was finally freed in 2003.

Lawyers call this type of mistake the prosecutor's fallacy, when people confuse the odds associated with a piece of evidence with the odds of guilt. Recognising this is also what eventually quashed the 1991 conviction for rape of Andrew Deen in Manchester. The courts realised at appeal that a one-in-three-million chance of a random DNA match for a semen stain from the crime scene did not mean there was only a one-in-three-million chance that anyone other than Deen could have been a match – those odds actually depend on the pool of potential suspects. In a population of 20 million adult men, for example, there could be as many as six other matches.

Now, Fenton and his colleague Amber Marks, a barrister and lecturer in evidence at Queen Mary, University of London, have begun assembling a group of statisticians, forensic scientists and lawyers to research a solution to bad statistics. "We want to do what people failed to do in the past, which is really get the legal profession and statisticians and probability guys understanding each other's language," says Fenton.

Their first job is to find out how often trials depend on Bayesian calculations, and the impact that the shoeprint-murder ruling might have on future trials. "This could affect thousands of cases," says Marks.

They have 37 members on their list so far, including John Wagstaff, legal adviser to the Criminal Cases Review Commission, and David Spiegelhalter, the Winton professor of the public understanding of risk at the University of Cambridge. Added to these are senior statisticians and legal scholars from the Netherlands, US and New Zealand.

Fenton believes that the potential for mathematics to improve the justice system is huge. "You could argue that virtually every case with circumstantial evidence is ripe for being improved by Bayesian arguments," he says.

But the real dilemma is finding a way to help people make sense of the calculations. The Royal Statistical Society already offers guidance for forensic scientists, to stop them making mistakes. Lagnado says that flowcharts in the style of family trees also help jurors visualise changing odds more clearly. But neither approach has been entirely successful. And until this complex bit of maths can be simply explained, chances are judges will keep rejecting it.

;