Appealing statistics

In the case of Regina *versus* Adams [1996], DNA evidence seemed to suggest that there was a 1 in 200 million chance that an innocent person would match the DNA found at the crime scene. **Peter Donnelly** explains how he subsequently became involved in the case and found himself trying to explain Bayes's Theorem to judge and jury.

The crime

The woman went out to a nightclub with friends, left, and walked home. Part of the walk was on a path along the side of a park, and as she walked through the park someone stopped her to ask the time. She was sure she'd had a good look at the man, and that—as she subsequently told the police—it was a face she would never forget. The man then attacked and raped her, and ran off. She reported the incident to the police, giving a fairly clear description of a man in his early 20s.

The police had no leads at the time, but quite some time later the woman thought she saw the man who had raped her walking along the street. She called the police, but by the time they arrived he had gone.

Some time later again, the police came up with a suspect, Denis Adams, and arranged an identity parade. The victim didn't pick him out. At the committal hearing, he was pointed out to her and she was asked if he matched her description, and she said no. She had described a man in his early 20s, and when asked how old Adams looked to her, she said around 40 (he was in fact 37).

However, there was very strong DNA evidence linking him with the crime and when the case came to trial in 1995, effectively the only incriminating evidence was that his DNA profile matched the DNA evidence found at the scene of the crime. The prosecution forensic scientist had calculated what is called a match probability, that is, the probability that if you pick someone at random, their DNA

would match the DNA sample of the assailant. That, according to him, was 1 in 200 million.

It's tempting for people not trained in statistics to get muddled, and to confuse two different probabilities. The first is the probability that a person would match the criminal's DNA profile given that they are innocent. The second is the probability that they are innocent given that they match the DNA profile. The forensic scientist's 1 in 200 million refers to the first probability. But jurors may wrongly think that this is the probability that the defendant is innocent. This misunderstanding is called the prosecutor's fallacy, and can be extremely prejudicial to the defendant. The error was often made in court by lawyers, and sometimes even by the forensic scientists, in the early days of DNA evidence. Who knows how often jurors make it in private?

But, even with the correct interpretation, the numbers are saying that the particular DNA profile is extremely rare. If the 1 in 200 million is correct, then it is probable that no one else in Britain has the same profile. Maybe there might be one other person, or just possibly two or even three.

> "This case was unusual in having DNA evidence pointing one way and all the other evidence pointing the other way"

The trial

This was the only incriminating evidence the jury heard. However, in Adams's defence there was the victim's own identification evidence; also, he had an alibi for the night involved, when his girlfriend said she spent the night with him. So there was a real issue here, for the lawyers in the first instance and ultimately for the jury, somehow to weigh up the prosecution evidence expressed in rather impressive hard numerical terms, against the more traditional defence evidence, expressed in terms juries are more familiar with—reported facts, alibis, and so on. I was asked by the defence barrister how one could compare the two.

I explained that, although difficult, there was a logically correct way: using Bayes's Theorem to reason with uncertainties by converting them into probabilities, and the statistical experts on the prosecution side agreed. The barrister then took the view that this should be explained to the jury, and it fell to me to do so in my evidence, in response to questions from counsel.

This case was quite unusual in having DNA evidence pointing one way and *all* the other evidence—which would often also be incriminating—pointing the other way. It was a really good example of the challenges faced in weighing up evidence—had it been invented as a hypothetical example, I suspect people would have thought it unrealistic! In other DNA cases, once there's a DNA match, there's normally another piece of evidence that incriminates, or at the very least there isn't a whole raft of other evidence that points towards innocence.

Here is an informal way of thinking about the evidence in this case. If the match probability were 1 in 1 million, then that means that, on average, there are 50 people in the UK who match. They may or may not have been in the right area on the night in question, and may or may not have been male and in the correct age range. However, the woman says that her attacker looked nothing like this particular suspect, who also happens to have an alibi, so it might well have been one of the other matching individuals who was the assailant, and a jury may acquit. If the number were 1 in 20 million, there are probably only a handful of people in the UK with that profile. They may be spread all over Britain, some may be women, or very young or very old, so a jury might think it was somewhat unlikely that another one of them was in the same place that night. If 1 in 200 million is right, it's unlikely, although still possible, that there is even one other person in Britain with the same DNA profile. In this case they may well think: if we *knew* there was no one else in Britain with that profile, we'd convict, and then just assume the victim misremembered what her assailant looked like and the girlfriend was lying or mistaken about the alibi.

So, unlike in many DNA cases, it really did matter whether the number was 1 in 2 or 20 or 200 million. The defence argued that for various reasons the number of 1 in 200 million figure was incorrect and, anyway, it is worth remembering that all of the statements about how many other matches there were in Britain exclude close relatives of the person on trial. As it happens, Adams did have a half-brother, whose DNA profile was unknown.

Adams was found guilty. Who knows what the jury thought—in this country you can't ask them. If you believe the 1 in 200 million number and you do all the Bayesian analysis, you might well convict. You might decide that, of the two explanations, the one that says the victim got muddled and the girlfriend was mistaken is the more plausible.

Appeal, retrial, appeal

There was an appeal, however, which was upheld because the Appeal Court felt that the judge hadn't instructed the jury clearly enough. He had told them that they could use Bayes's Theorem or not, as they wished. The Appeal Court said that the judge should have offered the jury more guidance on what to do if they didn't want to use Bayes's Theorem. In passing, they were very negative about the presentation of Bayes's Theorem in court, but, because this was not a matter on which they specifically needed to rule, their comments were not binding for future cases.

The Court of Appeal ordered a retrial, and in the retrial the defence said they wanted to present Bayes's Theorem to the jury. The judge agreed despite the Appeal Court comments.

In the retrial, there was again no dispute among the statistical experts from both sides that in principle Bayes's Theorem was the logically correct way to combine these sorts of evidence. Unlike at the first trial, though, the judge at the retrial asked the experts on both sides to get together and prepare a questionnaire which would help the jury to implement Bayes's Theorem, should they choose to do so. The two prosecution statisticians were from the Forensic Science Service. I knew them well professionally, and we collaborated on writing the questionnaire.

To apply Bayes's Theorem, you need to produce likelihood ratios for various different pieces of evidence. The

jury needed to ask themselves a series of questions such as: If he were the attacker, what's the chance that the victim would say her attacker didn't look anything like him? Also: If he *wasn't* her attacker, what's the chance she would say this? As you go through this exercise there are issues about how you agglomerate or separate various pieces of evidence. There's a lot of discussion about this in the legal literature on evidence. Provided one acts rationally (*coherently* is the technical term) the level of agglomeration will not affect the final answer. But that doesn't mean that real people necessarily reason the right way.

We were given instructions by the court about the level of agglomeration, and the jury ended up with about a dozen questions. The questionnaires were produced, and there were boxes where they could enter their numerical assessments, with a formula explaining how to combine them. The jury were told that this was in the experts' view the right way to do the reasoning, but that they were the jury and it was entirely up to them; they didn't have to use the questionnaires if they chose not to.

The episode had some amusing sidelines. It was suggested that it would be helpful to supply the jury (and judge) with basic calculators. Although the total cost was well under £100, this request was so unusual that it seemed to require clearance personally from the Lord Chancellor. Then, during my evidence, we walked the jury through a numerical example—the barrister would suggest token numbers in answer to the questions, and the jury and I entered them in the calculators which were eventually supplied. They seemed to have no difficulty in following this, but at an early stage in the calculation, when I said something to the effect that: "Your calculator should now show the value 31.6," and the jurors all nodded, the judge rather plaintively said: "But mine shows zero."

And, on separate occasions, first the defence barrister and then the learned judge corrected what they saw as a misprint on the questionnaire where we said that "the chance of throwing a six on a fair die" was about 17%. I carefully explained to each that "dice" was plural, and "die" was singular, and that we tended to be very careful about our professional use of language, much to the amusement, or perhaps bemusement, of the jury.

Again Adams was convicted, again there was an appeal—this time unsuccessful—and this time the court did rule on the use of Bayes's Theorem. They came down very heavily against it.

Not by means of a formula

I feel the judgment of the Court of Appeal oversimplified a very complex issue. They said that juries had been weighing up evidence for hundreds of years and that there was no reason to believe they couldn't do it now. They said that the jury should "evaluate evidence and reach a conclusion not by means of a formula, mathematical or otherwise, but by the joint application of their individual common sense and knowledge of the world to the evidence before them".

In a case such as Adams, there is a very live issue. The jury *somehow* have to weigh up the number of 1 in 200 million against ordinary evidence. Maybe, if the number is right, a rational juror would convict; if the number is 1 in 2 million, as the defence argued it might be, a rational person may well acquit.

> "It is only since the advent of DNA profiling that jurors have had to comprehend probabilities like 1 in 200 million"

There are right and wrong ways of reasoning with uncertainty, and there is plenty of documentary evidence that, particularly where small probabilities are concerned, people make mistakes when left to their own devices. So I think it's quite hard to argue in principle that juries should be denied the possibility of having experts explain to them the rules of logic that apply in this sort of situation. After all, juries hear expert witnesses in all sorts of settings: how you interpret financial accounts in fraud cases, how you interpret tyre skid marks in a road accident case, and so on. And having an expert explain the logic of reasoning with uncertainty certainly does not reduce the judicial process to a mathematical formula.

Secondly, it was a little disingenuous for the Appeal Court to say that juries have been considering these sorts of problems for hundreds of years. It's only since the advent of DNA profiling that jurors have had to comprehend probabilities like 1 in 200 million. It's not something that is in juries' experience collectively, nor in most people's experience. It is new and it throws up new problems.

So my own view is that it is troubling, as a matter of principle, to exclude evidence explaining how to reason with uncertainties. In the light of extensive evidence about people's routine errors in this setting, it seems inappropriate to deny a defendant the chance to have it explained carefully.

However, the pragmatic question is a very different one. Given the obvious constraints of the courtroom as a forum for teaching (no feedback, no homework, no exams...) one can and should ask whether the quality of judicial decision making would be improved or made worse by subjecting juries to explanations of Bayes's Theorem. Even though juries have little experience of thinking about numbers like 1 in 200 million, trying to explain how they should think may well end up making matters worse.

I should be clear that it was neither my suggestion nor my choice to explain Bayes's Theorem to the jury in the Adams case, and I remain unconvinced that it is a practicable way forward. In Adams, it was the defence barrister's choice, presumably because he felt it would increase the possibility of a favourable verdict for his client.

The Court of Appeal has subsequently provided some guidelines for similar situations. It advocated that judges should summarise cases in the following way. Suppose the match probability is 1 in 20 million. If you believe that number, then on average there will be 2 or 3 people in Britain whose DNA it could be, and probably no more than 6 or 7. That is the effect of DNA evidence: it narrows down the pool of possible suspects from everyone in the UK to 6 or 7 people—plus the accused. Now your job, as a member of the jury, is to decide, on the basis of the other evidence, whether or not you are satisfied that it is the person on trial who was the assailant, rather than one of the few other possible people who match. We don't know anything about the other people who match, although they are probably spread all over the UK, may have been nowhere near the scene of the crime, and some or all may also be ruled out by other factors, for example, gender or age.

I think this way of putting the DNA evidence is extremely helpful. It immediately steers the jury away from the error of the Prosecutor's Fallacy, and it is logically correct. In fact, it amounts to applying Bayes's Theorem to the DNA evidence *before* considering any of the other evidence. In this case a small prior probability of guilt would be appropriate, perhaps the inverse of the UK population. The effect of the evidence that the defendant's DNA matches the criminal's DNA is dramatically to increase the probability of guilt, to 1 in 5 or 6 or 7 in my example above. The juror's job is then to allow for the other evidence, to see whether or not they are convinced beyond reasonable doubt that the defendant is guilty.

But although, ironically given the earlier judgement, this suggestion by the Court of Appeal follows exactly the logic of Bayes's Theorem, it does not require that the Theorem be explained, and it moves the jury back into realms of reasoning with which they are more familiar. It doesn't solve all the problems, though. It becomes more difficult to put into practice when the match probability is 1 in 1 billion, or in more complex settings where the DNA profile is partially degraded, or involves a mixture from several individuals.

Peter Donnelly is a Professor of Statistical Science at the University of Oxford. He is interested in applications of probability and statistics in genetics, as well as statistical issues in DNA profiling.

Where has statistics taken you? If you would like to tell your personal statistical story, e-mail significance@rss.org.uk.