Although the concerns of Statistics and the Law might seem to have little to do with one other, they do share some fundamental common interests, such as interpretation of evidence, hypothesis testing, and decision-making under uncertainty. Philip Dawid takes up questions that, though long of interest to a few scholars, have only recently become the focus of much attention.

Many pioneers of probability and statistics, including the Bernoullis, Condorcet, Laplace, Poisson and Cournot, were motivated by problems of quantification and combination of legal evidence and judgement. But the trail they blazed became disused and overgrown as statisticians lost interest in such questions while legal evidence scholars confined themselves to issues of admissibility, precedent and other such formal rules, paying remarkably little attention to problems of interpretation. Very occasionally some aspect of statistical evidence or argument would break surface—significant legal cases include the 1865 Howland will case and the 1894 Dreyfus case—but it was not until the 1968 Californian case of People v. Collins, in which the prosecution presented a fallacious statistical argument in an attempt to magnify the impact of eye-witness identification evidence, that the issues became subject to serious discussion and argument. This stimulated what became known in the academic legal community as “the new evidence scholarship”, with emphasis on the content and meaning of evidence, and on the “rationalist” principles to which it should be subject—including a lively debate about the role of probability theory in the interpretation and management of evidence of all kinds.

The probabilities debate remained academic until 1985 when, with the advent of DNA profiling and its numerical “random match probabilities”, presentation of, and argument about, statistical evidence of various kinds started to become much more common in the courts. Although there were, and still are, many cases where such evidence has been shockingly misused, the adversarial system has ensured that the underlying logic has not been without scrutiny. In particular, much excellent logical and statistical analysis has been developed by the UK Forensic Science Service, to identify and eliminate potential weak spots in its handling of DNA evidence.

A conflict between approaches
The application of probability in cases at law presents some special difficulties, beyond mere unfamiliarity, for both lawyers and statisticians. Some of these
stem from an internal tension in the law itself between the resolution of individual cases and wider issues of justice and fairness. A much discussed toy example due to Jonathan Cohen involves a rodeo where some spectators climbed over the fence to avoid paying. The impresario sues a spectator for evading the entrance charge; there is no relevant evidence except that 499 people paid and 1000 people were counted on the seats. Does the case deserve to succeed “on the balance of probabilities” (the appropriate civil standard)? It has been argued that to find for the plaintiff would mean, unfairly, that he could similarly succeed against every one of the spectators; while to reject his claim would mean, even more unfairly, that he would have no redress at all against the cheats. But this is not a fault of probability—identical problems would afflict any consistently applied method of reaching a firm conclusion. The difficulty is a legal one, arising from the procedural requirement that once a case is brought to law a definitive verdict must be reached, however inconclusive the evidence.

A more serious concern is whether, to be admitted as relevant, a probability should relate directly to the specific case at hand, rather than (as for the rodeo) merely describing non-specific background knowledge. While a statistician would, as a matter of logic, be unhappy to make such a distinction, a lawyer might regard it as natural and desirable on procedural grounds. Again probability is not of the essence here: the conflict is between rationalist and procedural approaches to evidence.

A related issue is the relevance of background information to the prior probability (i.e., before any specific evidence has been introduced) that the accused is guilty as charged. The popular Bayesian statistical approach to the analysis of evidence requires that this be assessed, so that it can be combined, by well-determined mathematical rules, with the case-specific evidence—finally leading to a posterior probability of guilt that takes all the evidence, both specific and non-specific, into account. But even when a prior probability can be regarded as reasonably objective, being based on population figures, it might be legally excluded on procedural grounds of non-specificity.

Exclusion of evidence of any kind seems to pose a problem for statisticians, who have theorems that tell them that more information is always better than less. But these relate to a single decision maker who has carefully and unambiguously quantified all his or her uncertainty. It is not hard to present rationalist arguments for the exclusion of some legal evidence—for example, where reasonable people might disagree markedly as to its relevance and force, or find it hard to assess appropriately. Thus evidence of similar prior convictions is not normally admissible, although it may be revealed, often to the dismay of the jury, after the verdict is in. This is reasonable in view of the fact that it would be all too easy for the jury to overestimate the impact of such evidence—forgetting that the police are more likely to pick up and put on trial a seasoned offender, whether he is guilty in this case or not. (This perfectly justifiable behaviour of the police should also be borne in mind when reading the habitual lurid press reports, after a
murder conviction, about the nasty habits and generally weak character of the accused—as if these were a definitive vindication of the verdict.) The Government is currently proposing substantial weakening of the rules against revealing prior convictions. If this happens then particular care will be needed in dealing with such evidence.

Other non-rationalist exclusions are fully justifiable and intelligible on broader policy grounds. Thus the law recognises a right not to incriminate oneself or one’s spouse, to the extent that no inference is to be drawn from refusal to testify. Any attempt at rationalistic analysis simply has to accept and live with such procedural exclusions. To echo Thomas More in Robert Bolt’s “A Man for All Seasons”: “The world must construe according to its wits. This Court must construe according to the law.”

What of the “presumption of innocence”? This can be interpreted, semi-quantitatively, as requiring that at the start of a trial a juror should not consider the accused as any more or less likely to be guilty than anyone else. Now it might seem that, on the contrary, a rationalist assessment would take into account that the police and Director of Public Prosecutions do not usually bring a case to court unless there is a high chance that the accused can be convicted. Surely this knowledge raises the prior probability of guilt? Again the legal rule appears to be excluding evidence that, on a rationalist analysis, appears relevant. However, this analysis is too simplistic. Some of the evidence leading to the decision to prosecute will itself be presented later in court, and it would be wrong to count it twice; while any such evidence that is not so presented should quite properly be ignored.

**Introducing background evidence**

But things are rarely clear-cut. In a trial for a rape in London, it is surely both rationally and procedurally correct to regard the male defendant as more likely to be guilty than any female. But should we regard his age, between 16 and 60, as making him more likely to be guilty than those outside this age group? Or his London address as making him a more likely perpetrator than a stay-at-home New Yorker? Many academic probabilistic analyses, and some presented in court, have taken such background demographic information as relevant, and used it, in more or less straightforward ways, to help in quantifying a suitable prior probability of guilt. Is such evidence admissible? Does its use violate the presumption of innocence?

To the contrary, arguments that start by identifying a suitable pool of potential perpetrators are almost invariably presented in court by defence counsel, in order to counter the “prosecutor’s fallacy”. An example is when the identical DNA profile, present in 1 in 1 million people, is obtained both from the defendant and from blood found at the crime-scene. For simplicity we here take it as certain that the blood belongs to the guilty party, although this would often be subject to dispute. The prosecutor now argues that 1 in 1 million is the probability that the
DNA match occurred by innocent means, and thus (entirely negligibly) that the defendant is innocent—an intuitively appealing, though thoroughly illogical, interpretation. But if defence counsel can justify a population of potential perpetrators of size 5 million, he can argue that there would be 5 innocent matches in addition to the guilty party. Since the DNA does not distinguish between the 6 who match, of which the defendant is one, the probability he is the guilty party is only 1 in 6.

The case of Sally Clark, initially convicted of murdering her two babies against a defence of death by natural causes, highlights other issues of background evidence. The prosecution led evidence of the incidence of cot deaths—SIDS—in families such as Sally’s. A (highly dubious) calculation gave an estimate of 1 in 73 million for the probability of two such deaths. But if background evidence of double death rates due to SIDS (or other natural causes) is relevant, then surely so too should be background evidence of double death rates due to murder. To present either of these figures without some assessment of the other would be both irrational and prejudicial.

Even were we to have reliable values for both figures, there would remain the problem of what to do with them. Prosecution and defence might each grab one of the figures and argue as to whose is tinier. But probabilistic analysis shows that the overall impact of this background evidence is entirely carried by the ratio of the two figures. It also shows exactly how this ratio should be combined with other evidence (there was also medical evidence in the Sally Clark case, and indeed it was this that led to her release at second appeal). Although none of the relevant calculations are at all deep or difficult, the logic that identifies the correct statistical approach can be surprisingly subtle, and those not well versed in it frequently make serious errors. Part of the problem lies with a view of Statistics as something any fool can do, as instanced by the refusal of the court at Sally Clark’s first appeal to take oral testimony from two statistical expert witnesses on the grounds that “It’s hardly rocket science, is it?” In fact, interpretation of statistical evidence can not be safely left in the hands of the uninitiated. Whether or not it is considered appropriate for statistical expert witnesses to instruct judge and jury in the correct logical handling of evidence, statisticians certainly have a role to play in improving the general understanding of these issues among lawyers.

Even—indeed, especially—when a rationalist analysis of evidence appears to conflict with procedural requirements and rules, it is still important to undertake it, if only to see just how far the law departs from its prescriptions: by such ways might changes for the better come about. And, where probabilistic and statistical evidence is admitted, it is essential that this be treated with the respect it deserves.
Philip Dawid, Professor of Statistics at University College London, has a long-standing interest in statistical arguments in law. As director of an interdisciplinary research programme on “Evidence, Inference and Enquiry”, he is also concerned with wider issues of the rational use of evidence.