LEGAL PROBABILISM

MARCELLO DI BELLO – ASU - SEPTEMBER 19, 2024

NOTES ON DAWID'S BAYES'S THEOREM AND WEIGHING EVIDENCE BY JURIES

1. Dawid's paper is a good overview of how evidence can be combined and weighed in court using probability theory, in particular Bayes's theorem.

2. The first thing to remember is the odds version of Bayes's theorem:

$$\frac{P(G|E)}{P(\overline{G}|E)} = \frac{P(E|G)}{P(E|\overline{G})} \times \frac{P(G)}{P(\overline{G})}.$$

   In other words,

$$\textit{posterior odds} = \textit{likelihood ratio} \times \textit{base rate odds}.$$

   Note that G and $\overline{G}$ are the two hypotheses that are being compared, say one is the guilt hypothesis by the prosecutor and the other is the innocence hypothesis by the defense. What these hypotheses say, exactly, is up to the parties in the litigation.

3. Each bit in the equation is important:

   (a) The posterior odds are not the same thing as the posterior probability. The posterior probability $P(G|E)$ is given by $\frac{PO}{1+PO}$, were PO are the posterior odds provided G and $\overline{G}$ are one the negation of the other, or else the posterior probability cannot be inferred from the posterior odds. Can you say why?

   (b) The likelihood ratio is a measure of the strength of the evidence for/against the two competing hypotheses.

   (c) The priors odds are also important. The lower the prior odds, the lower the posterior odds (the lower the posterior probability), with the same likelihood ratio. See table 1 on page 4.

4. Strictly speaking we always start with some information or evidence, however generic. So a better formula would be:

$$\frac{P(G|E \ \& \ B)}{P(\overline{G}|E \ \& \ B)} = \frac{P(E|G \ \& \ B)}{P(E|\overline{G} \ \& \ B)} \times \frac{P(G|B)}{P(\overline{G}|B)},$$

   where B is the background information we always start out with. This notation, however, can be confusing and perhaps overly complicated. So, for the sake of simplicity, it is often avoided. We do that in all the examples below.

5. The odds version of Bayes' theorem can be applied to the Sally Clark case:

$$\frac{P(M|E)}{P(S|E)} = \frac{P(E|M)}{P(E|S)} \times \frac{P(M)}{P(S)},$$

   where M is the murder hypothesis (Sally Clark murdered her babies) and S is the SIDS hypothesis (the babies died of cot death). What is E? The fact that the two babies were found dead.

6. Next, plug in some numbers:
   (a) $\frac{P(E|M)}{P(E|S)} = 1$. (Why?)
   (b) $\frac{P(M)}{P(S)}$ = 1/8.4 billion/1/73 million $\approx$ 0.0009. (Why?)

   So the posterior odds are 9/1000, or in other words above 100:1 against M.

7. We could have picked different hypotheses, for example, the murder hypothesis M, and as alternative the negation of $\neg M$ (that Sally did not murder her babies). So:

$$\frac{P(M|E)}{P(\neg M|E)} = \frac{P(E|M)}{P(E|\neg M)} \times \frac{P(M)}{P(\neg M)},$$

8. Next, plug in some numbers:
   (a) $\frac{P(E|M)}{P(E|\neg M)}$ =1/(1/73)=73 million. (Why?)
   (b) $\frac{P(M)}{P(\neg M)} \approx$ 1/8.4 billion. (Why?)

   So the posterior odds are again 9/1000, or in other words above 100:1 against M.

9. The posterior odds are the same, but note that the likelihood ratios are very different: 1 in the earlier analysis and 73 million in the second analysis. It looks as though the same evidence E can be both irrelevant and extremely probative of guilt. How do we explain this shift?

10. So, convict or not convict? This very much depends on the decision criterion we follow.
    (a) Option 1: Since the posterior probability of guilt is low, do not convict.
    (b) Option 2: Since the likelihood ratio is so high (second analysis above), convict.
    (c) Option 3: Under Neyman-Pearson approach (2.2.2), convict. More specifically:
        i. If Sally Clark murder her babies, to convict is the correct decision. No error.
        ii. If the babies died of natural causes, to convict is the incorrect decision, but this error would be committed only 1 in 73 million, an extremely unlikely event.
        iii. By convicting, either no error is committed or an error is extremely unlikely.

11. The odds version of Bayes's theorem can also be used to analyze match identification evidence, call it M. The methodology is the same:

$$\frac{P(G|M)}{P(\overline{G}|M)} = \frac{P(M|G)}{P(M|\overline{G})} \times \frac{P(G)}{P(\overline{M})},$$

    where we set $P(M|G) = 1$ and $P(M|\overline{G}) =$*random match probability*. (Why?)

12. Dawid looks at two contradictory informal arguments about match identification evidence, the prosecution argument (3.1.1) and the defense argument (3.1.2). What are these arguments and how does Bayes's theorem help to see which one is correct?

13. Let's now complicate things further. What if there is more than one item of evidence in a case, for example, some evidence favors the prosecution hypothesis and other evidence favors the defense hypothesis? The evidence in the R v Adams case is a good example.

   (a) This was a rape case. They found a match between the defendant, Adams, and the crime scene sample. Random match probability was 1 in 200 million, though the defense said it could be as low as 1 in 2 million. We need to account for this variability. The match is evidence against the defendant.

   (b) The other evidence was in favor of the defendant. The victim did not pick Adams out of a lineup. Also, the victim said Adams did not resemble the attacker. There was also alibi evidence. Adams's girlfriend said he was with her when the crime took place. How do we put everything together?

14. How do we put everything together? Let's proceed piece by piece:

   (a) Prior odds can be set to 1 in 200,000 because 150,000 adult males could have committed the crime in the area. This number can go up to 200,000 to be conservative. At this point we assume we don't know anything else, just where the crime was committed and that Adams lived in the vicinity.

   (b) The victim did not recognize Adams: the likelihood ratio of this piece of evidence can be set to 1/9. (Why?)

   (c) Alibi evidence: likelihood ratio can be set to 1/2. (Why?)

   (d) Match identification evidence: likelihood ratio between 2 million and 200 million.

15. Put everything together by multiplying each ratio (with upper and lower bounds):

$$\textit{prior odds} \times \textit{LR lineup} \times \textit{LR alibi} \times \textit{LR match} = \textit{posterior odds}.$$

$$1/200,000 \times 1/9 \times 1/2 \times 2,000,000 \approx 0.55$$

$$1/200,000 \times 1/9 \times 1/2 \times 200,000,000 \approx 55$$

   These are posterior ratios, not posterior probabilities of guilt. Dawid says that the posterior probability of guilt would be between 36% and 98%. How does he get these figures?

16. A couple of things are notable about this case:

   (a) The jury ended up convicting Adams; the decision was appealed; there was a retrial and the jury convicted again; the second appeal was dismissed. So the jury deciding against what Bayes's theorem seems to recommend. Was the probability of guilty high enough since it could be anywhere between 36% and 98%?

   (b) In both trials, the jury was instructed to use Bayes's theorem to weigh the evidence, but the appellate court stated that the Bayesian approach should not be used in court.

17. A peculiar type of case is one in which a suspect is identified through a DNA database search, called a cold-hit case. Does the value of the DNA match evidence differ in cold-hit cases? On this question, Dawid compares the prosecution argument (4.1) and the defense argument (4.2). Which one is correct given the Bayesian approach?