

Elsevier required licence: © <2018>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Accepted Manuscript

Title: Bayesian interpretation of discrete class characteristics

Author: Dennis McNevin

PII: S0379-0738(18)30766-7

DOI: <https://doi.org/10.1016/j.forsciint.2018.09.013>

Reference: FSI 9482

To appear in: *FSI*

Received date: 1-3-2018

Revised date: 2-9-2018

Accepted date: 15-9-2018

Please cite this article as: Dennis McNevin, Bayesian interpretation of discrete class characteristics, Forensic Science International <https://doi.org/10.1016/j.forsciint.2018.09.013>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Bayesian interpretation of discrete class characteristics

Dennis McNevin^{a,b,*}

^a Centre for Forensic Science

School of Mathematical and Physical Sciences (MaPS)

Faculty of Science

University of Technology Sydney

Broadway, NSW, 2007

Australia

^b National Centre for Forensic Studies

Faculty of Science and Technology

University of Canberra

Bruce, ACT, 2617

Australia

* Corresponding author: dennis.mcnevin@uts.edu.au

Highlights

- Bayesian interpretation of single discrete class characteristics is explored.
- Posterior odds ratios favourable to the prosecution cannot be attained.
- This can be overcome by analysing more than one class characteristic.
- We define a criterion to determine how many characteristics are required.

- A frequentist interpretation would not reveal this limitation.

Abstract

Bayesian interpretation of forensic evidence has become dominated by the likelihood ratio (LR) with a large LR generally considered favourable to the prosecution hypothesis, H_P , over the defence hypothesis, H_D . However, the LR simply quantifies by how much the prior odds ratio of the probability of H_P relative to H_D has been improved by the forensic evidence to provide a posterior ratio. Because the prior ratio is mostly neglected, the posterior ratio is largely unknown, regardless of the LR used to improve it. In fact, we show that the posterior ratio will only favour H_D when LR is at least as large as the number of things that could possibly be the source of that evidence, all being equally able to contribute. This restriction severely limits the value of evidence to the prosecution when only a single, discrete class characteristic is used to match a subset of these things to the evidence. The limitation can be overcome by examining more than one individual characteristic, as long as they are independent of each other, as they are for the genotypes at multiple loci combined for DNA evidence. We present a criterion for determining how many such characteristics are required. Finally, we conclude that a frequentist interpretation is inappropriate as a measure of the strength of forensic evidence precisely because it only estimates the denominator of the LR.

Keywords: Bayes; Frequentist, Prior, Posterior, Likelihood ratio, Forensic evidence

Novelty

This manuscript demonstrates a limitation of the forensic interpretation of discrete class characteristics which is only apparent from a Bayesian interpretation. As such, it demonstrates the weakness of a frequentist interpretation, as well as providing guidance for the minimum requirements for analysis of class characteristics for evidence to be favourable to the prosecution.

1. Introduction

Bayes theorem is used to measure the relative strengths of two mutually exclusive hypotheses as a result of evidence presented to the court [1-6]. These are generally the prosecution and defence hypotheses (H_P and H_D , respectively). The prosecution hypothesis often asserts that someone or something (a tool, a weapon, a car, for example) is the source of evidence found at a crime scene. The corresponding defence hypothesis asserts that this person or thing is not the source of the evidence. A Bayesian interpretation of these two competing hypotheses involves adjusting belief in their relative truths based on observing how often there have been instances that support them, or as Richard Price noted in the forward to the posthumous publication of Bayes theorem [7]:

Common sense is indeed sufficient to show us that, from the observation of what has in former instances been the consequence of a certain cause or action, one may make a judgement what is likely to be the consequence of it another time and that the larger number of experiments we have to support a conclusion, so much more the reason we have to take it for granted.

It is illustrative to demonstrate the concept using the analogy of a deck of cards. The individual cards (three of hearts, nine of diamonds, king of clubs, *etc*) are analogous to the people or things that could

possibly be the source of the evidence. The cards have class characteristics that define a subset of the complete deck. For example, spades constitute one quarter ($13/52$) of the deck (neglecting jokers) and aces constitute one thirteenth ($4/52$) of the deck. Hence knowing a class characteristic of a card narrows the pool of possible cards. This has a direct forensic analogy in that, for example, knowing the blood group type of an evidentiary blood sample narrows the pool of potential donors and knowing the colour of a paint chip narrows the pool of potential cars it could be derived from.

Let us then consider a card randomly drawn from a deck. The probability that this card is a spade (S) is given by:

$$P(S) = \frac{13}{52} = \frac{1}{4}$$

The probability that the card is not a spade (the alternative hypothesis) is given by:

$$P(\bar{S}) = \frac{39}{52} = \frac{3}{4} = 1 - P(S)$$

It follows that for an ace (A) and an ace of spades (AS):

$$P(A) = \frac{4}{52} = \frac{1}{13} \text{ and } P(\bar{A}) = \frac{48}{52} = \frac{12}{13}$$

$$P(AS) = \frac{1}{52} \text{ and } P(\overline{AS}) = \frac{51}{52}$$

If $P(AS)$ is the prior or original probability that the card is an ace of spades, then we can update this probability if we know that the card is a spade. The posterior probabilities of AS and \overline{AS} are then:

$$P(AS/S) = \frac{1}{13} \quad \text{and} \quad P(\overline{AS}/S) = \frac{12}{13}$$

Hence we have a higher probability that the card is an ace of spades (and a lower probability that it is not) given that we know it is a spade. If we know that the card is an ace, we can further increase the posterior probability of an ace of spades:

$$P(AS/A) = \frac{1}{4} \quad \text{and} \quad P(\overline{AS}/A) = \frac{3}{4}$$

This situation is analogous to having a better forensic test. For example, we can achieve higher discrimination between people (further narrow the potential donors) if we use DNA profiling rather than blood group typing.

For our card analogy, Bayes theorem tells us that:

$$P(AS/A) \times P(A) = P(A/AS) \times P(AS) \quad \text{or} \quad \frac{1}{4} \times \frac{1}{13} = 1 \times \frac{1}{52} = \frac{1}{52}$$

$$P(\overline{AS}/A) \times P(A) = P(A/\overline{AS}) \times P(\overline{AS}) \quad \text{or} \quad \frac{3}{4} \times \frac{1}{13} = \frac{3}{51} \times \frac{51}{52} = \frac{3}{52}$$

If we divide these two equalities:

$$\frac{P(AS/A) \times P(A)}{P(\overline{AS}/A) \times P(A)} = \frac{P(A/AS) \times P(AS)}{P(A/\overline{AS}) \times P(\overline{AS})}$$

This leads to:

$$\frac{P(AS/A)}{P(\overline{AS}/A)} = \frac{P(A/AS)}{P(A/\overline{AS})} \times \frac{P(AS)}{P(\overline{AS})}$$

This is the form of the familiar expression for Bayesian inference in a forensic setting [8] where evidence, E , is presented to the court and H_P and H_D are mutually exclusive hypotheses, *ie.* $P(H_P) = 1 - P(H_D)$:

$$\frac{P(H_P/E)}{P(H_D/E)} = \frac{P(E/H_P)}{P(E/H_D)} \times \frac{P(H_P)}{P(H_D)} = \text{LR} \times \frac{P(H_P)}{P(H_D)}$$

The likelihood ratio (LR) or “Bayes factor” [9-11] is the ratio of the probabilities of observing the evidence given the prosecution and defence hypotheses. In our card analogy, the LR is dependent on the class characteristic known about the questioned card: the suit (spade, club, diamond or heart) or the rank (A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K). We develop this analogy to demonstrate the limited value of a single class characteristic in providing a posterior ratio favourable to the prosecution.

2. Methods

Continuing our card analogy, the LR is the ratio of the probabilities of observing a spade (or an ace) given the alternate hypotheses that the card is or is not the ace of spades. In the case that we know the card is a spade:

$$\frac{P(AS/S)}{P(\overline{AS}/S)} = \frac{P(S/AS)}{P(S/\overline{AS})} \times \frac{P(AS)}{P(\overline{AS})}$$

Substituting numerical probabilities:

$$\frac{1/13}{12/13} = \frac{1}{12/51} \times \frac{1/52}{51/52}$$

$$\frac{1}{12} = 4.25 \times \frac{1}{51}$$

It is 4.25 times more likely that the card is a spade under the hypothesis that it is the ace of spades than under the hypothesis that it is not the ace of spades. Note that we cannot say that it is 4.25 times more likely than not that the card is an ace of spades given that is a spade. It is in fact 12 times less likely. But this is still an improvement on the prior ratio (1/51). For the same reason, we cannot say in court that the prosecution hypothesis is LR times more likely than the defence hypothesis given the evidence. This is the well documented fallacy of the transposed conditional or “prosecutor’s fallacy” [5, 12, 13]. We can only say that the evidence is LR times more likely under the prosecution hypothesis than under the defence hypothesis.

In the case that we know the card is an ace:

$$\frac{P(AS/A)}{P(\overline{AS}/A)} = \frac{P(A/AS)}{P(A/\overline{AS})} \times \frac{P(AS)}{P(\overline{AS})}$$

$$\frac{1/4}{3/4} = \frac{1}{3/51} \times \frac{1/52}{51/52}$$

$$\frac{1}{3} = 17 \times \frac{1}{51}$$

Both the LR and hence the posterior ratio have increased because our new class characteristic (rank = ace) is more discriminating than our old one (suit = spade). However, the posterior ratio is still less than one. In other words, it is still less likely that the unknown card is an ace of spades than it is not an ace of spades, in spite of a significantly larger LR. However, in a forensic setting, it is the posterior ratio that is of ultimate interest to the court. The problem is that it is never reported because the prior ratio is not generally known. Under what conditions, then, is the posterior ratio likely to favour the prosecution?

3. Results

Let N be the total number of things (people, tools, weapons, cars, *etc*) equally capable of being the source of crime scene evidence. Let n be the number of things (a subset of N) that have the same class characteristic as the evidence. We assume that each of the things is equally likely to have contributed to the evidence and so the prior probabilities of H_P and H_D are:

$$P(H_p) = \frac{1}{N} \text{ and } P(H_d) = \frac{N-1}{N}$$

The probability of observing the evidence given the prosecution hypothesis is:

$$P(E/H_p) = 1$$

The probability of observing another thing with the same class characteristic as that observed in the evidence (the defence hypothesis) is then:

$$P(E/H_d) = \frac{n-1}{N-1}$$

Bayes theorem then tells us:

$$\frac{P(H_p/E)}{P(H_d/E)} = \frac{P(E/H_p)}{P(E/H_d)} \times \frac{P(H_p)}{P(H_d)} = \frac{1}{(n-1)/(N-1)} \times \frac{1/N}{(N-1)/N} = \frac{N-1}{n-1} \times \frac{1}{N-1} = \frac{1}{n-1}$$

Hence, the prior ratio is $1/(N-1)$, the LR is $(N-1)/(n-1)$ and the posterior ratio is always $1/(n-1)$.

Table 1 demonstrates that the posterior ratio will only favour the prosecution hypothesis when $n < 2$.

In other words, of all the possible things that could be the source of the evidence (numbering N), if any more than one of them have ever been observed to have the same discrete class characteristic as the evidence, the posterior ratio will never be more than one. For our card analogy, the probability of drawing an ace of spades will never be greater than the probability of not drawing an ace of spades

unless we know the card is an ace of spades. This means that, in order to produce a posterior ratio favourable to the prosecution in court, the frequency of the discrete class characteristic, f , must be less than $2/N$.

We can express the *posterior* ratio in terms of the LR:

$$\frac{P(H_p/E)}{P(H_d/E)} = \text{LR} \times \frac{P(H_p)}{P(H_d)} = \text{LR} \times \frac{1}{N-1}$$

This relationship is shown in Figure 1 where it can be seen that the posterior ratio favours the prosecution hypothesis (is greater than one) for small N and large LR and, conversely, favours the defence hypothesis (is less than one) for large N and small LR. The posterior ratio is infinitely large (solely favours the prosecution hypothesis) for $N = 1$ as, in this case, there is only one possible person or thing that could be the source of the evidence.

We can determine what values for LR will provide a posterior ratio favourable to the prosecution:

$$\text{LR} = \frac{N-1}{n-1}$$

Rearranging and solving for n :

$$n = \frac{N-1}{LR} + 1 < 2$$

Therefore:

$$\frac{N-1}{LR} < 1 \text{ and } LR > N-1$$

Hence the LR must be at least as large as the number of things that could possibly have been the source of the evidence, N . For our card analogy, we would need $LR > 51$ for the posterior ratio to favour drawing an ace of spades and this would only be the case if we knew the card was an ace of spades.

Given that a single class characteristic will never produce a posterior ratio favourable to the prosecution for $n > 1$, we need to know how many independent class characteristics are required to provide a posterior ratio favourable for the prosecution under these conditions. Let the number of things with the same class characteristics be $n_1, n_2, n_3, \dots, n_k$, where each of the k characteristics is independent of the others. In order to produce a posterior ratio greater than one:

$$\frac{2}{N} > \frac{n_1}{N} \frac{n_2}{N} \frac{n_3}{N} \dots \frac{n_k}{N} = f_1 f_2 f_3 \dots f_k$$

where f_i is the frequency of the i th class characteristic. Taking the logarithm base 2 of both sides (for convenience):

$$\ln_2 2 - \ln_2 N > \ln_2 f_1 + \ln_2 f_2 + \ln_2 f_3 \dots$$

$$\ln_2 N < 1 - \ln_2 f_1 - \ln_2 f_2 - \ln_2 f_3 \dots$$

$$2^{1 - \ln_2 f_1 - \ln_2 f_2 - \ln_2 f_3 \dots} > N$$

We now have a criterion for determining how many individual class characteristics are required to produce a posterior ratio favourable for the prosecution if we know the frequencies of those class characteristics and the number of things that could equally have been the source of the evidence. So, for example, consider a paint chip that could equally have been derived from 1000 cars. It has a colour (pink) that is present in 1 % of cars.

$$2^{1 - \ln_2 0.01} = 2^{1 + 6.64} = 2^{7.64} = 200 < N$$

As we have already demonstrated, a single class criterion is insufficient to produce a posterior ratio greater than one. We therefore need another one. Let's assume that the paint chip has a solvent residue that is present in 5 % of cars, independent of paint colour. We now have a second class characteristic and:

$$2^{1 - \ln_2 0.01 - \ln_2 0.05} = 2^{1 + 6.64 + 4.32} = 2^{11.97} = 4,000 > N$$

We have achieved our criterion and the combination of these two independent class characteristics will produce a posterior ratio favourable to the prosecution. Adding a third class characteristic will produce an even greater LR and posterior ratio. The magnitude of the posterior ratio produced is:

$$\begin{aligned} \frac{P(H_p/E)}{P(H_D/E)} &= \frac{P(E/H_p)}{P(E/H_D)} \times \frac{P(H_p)}{P(H_D)} = \frac{1}{\frac{n_1-1}{N-1} \frac{n_2-1}{N-1} \frac{n_3-1}{N-1} \dots \frac{n_k-1}{N-1}} \times \frac{1/N}{(N-1)/N} \\ &= \frac{(N-1)^k}{\prod_i (n_i-1)} \times \frac{1}{N-1} = \frac{(N-1)^{k-1}}{\prod_i (n_i-1)} \end{aligned}$$

For large N :

$$\frac{P(H_p/E)}{P(H_D/E)} \approx \frac{N^{k-1}}{\prod_i (n_i-1)}$$

For large n :

$$\frac{P(H_p/E)}{P(H_D/E)} \approx \frac{N^{k-1}}{\prod_i n_i} = \frac{N}{N} \frac{N^{k-1}}{\prod_i n_i} = \frac{1}{N} \frac{N^k}{\prod_i n_i} = \frac{1}{N} \frac{1}{\prod_i f_i}$$

So, for example, consider again the paint chip with colour (pink) that is present in 1 % of 1000 cars (*ie.* 10 cars):

$$\frac{P(H_p/E)}{P(H_D/E)} = \frac{(N-1)^{k-1}}{\prod_i (n_i-1)} = \frac{1}{10-1} = \frac{1}{9}$$

This posterior ratio is favourable to the defence, not the prosecution. There are nine other cars that could have been the source of the paint chip. Consider again the solvent residue that is present in 5 % of cars (*ie.* 50 cars). With this second class characteristic:

$$\frac{P(H_p/E)}{P(H_D/E)} = \frac{(1000-1)^{2-1}}{(10-1)(50-1)} = \frac{999}{9 \times 49} = \frac{111}{49} \approx 2.27$$

This posterior ratio is now favourable to the prosecution.

4. Discussion

We often have to make conservatively large estimations of N because we don't know how many things could have possibly been the source of the evidence. For example, we may need to include all the people or cars or weapons in a country or in the world as possible sources. Under these types of assumptions, $f < 2/N$ is rarely the case and so the posterior probability is unlikely to favour the prosecution if a single discrete class characteristic is analysed.

We can circumvent this problem by analysing more than one independent class characteristic. This has most successfully been achieved for DNA evidence where multiple class characteristics

(genotypes at multiple loci) are compared with frequencies of the same characteristics in the general population. Population allele frequencies are used to infer genotype frequencies at multiple, independent loci such that the individual genotype frequencies at each locus can be multiplied together to produce overall genotype frequencies that are much less than $2/N$ (the so called product rule) [14]. However, this same principle should also be applied to other forms of evidence. It is unlikely that determining the frequency of pink cars is going to provide a posterior ratio favourable to the prosecution in a case where the evidence consists of a pink paint chip, no matter how low the frequency (and thus, how high the LR). This is because we need only observe more than one pink car to make the frequency $2/N$ or greater. We need to analyse a further independent (of colour) class characteristic in order to be able to reduce the combined frequency. This may be the presence of a particular mineral or solvent in the paint (but not pigment, as this is likely to contribute to colour).

We have developed a criterion for the number of independent class characteristics required to produce evidence favourable to the prosecution (*ie.* to produce a posterior ratio greater than one. For N things equally likely to have been the source of evidence and k independent characteristics of those things with frequencies $f_1, f_2, f_3, \dots, f_k$, a posterior ratio greater than one will result when:

$$2^{1 - \ln_2 f_1 - \ln_2 f_2 - \ln_2 f_3 - \dots - \ln_2 f_k} > N$$

For large enough N , the posterior ratio will be:

$$\frac{P(H_p|E)}{P(H_d|E)} = \frac{1}{N} \frac{1}{\prod_i f_i}$$

This is a result familiar to forensic biologists who apply the product rule for genotype frequencies at different (independent) DNA loci and assume that there are N individuals who are equally likely to have been the source of that DNA (before knowledge of the DNA profile).

It is easy to be seduced by large LR_s. Modern short tandem repeat (STR) DNA profiling kits produce LR_s greater than 10^{20} for matching DNA profiles. This is only meaningful, however, if we can estimate the prior probability that a random member of the population was the DNA donor [15]. This requires knowledge of the number of people that could possibly have had access to a crime scene at a particular time (N). If we do not know this then we must necessarily conservatively overestimate N as, for example, the number of people in the region, country or, most conservatively, the world. It is at least conceivable that anyone in the modern world could travel to the crime scene, deposit their DNA and then travel away again. At the time of writing, the world population was estimated at 7.4 billion so that, in the worst case (for the prosecution):

$$\frac{P(H_p|E)}{P(H_d|E)} = LR \times \frac{1}{N-1} \approx \frac{LR}{N} = \frac{10^{20}}{7.4 \times 10^9} = 1.4 \times 10^{10}$$

This posterior ratio still represents very strong evidence for the prosecution but consider a partial DNA profile with $LR = 10^8$. At first glance, this would still seem strong evidence for the prosecution but, under the assumption that anyone in the world could be the donor:

$$\frac{P(H_p|E)}{P(H_d|E)} = \frac{LR}{N} = \frac{10^8}{7.4 \times 10^9} = 0.01$$

Hence, conservatively, the defence hypothesis is 100 times more likely given this partial profile match. While LR_s of these magnitudes are common for DNA evidence, they are rare (or absent) for other forms of evidence.

A frequentist interpretation of forensic evidence only accounts for the denominator in the LR and takes no account of the prior and thus the posterior probability ratios. The frequentist only considers the blue line in Figure 1 and not the red line but it is the posterior ratio relative to the red line that is the concern of the court. Because we have shown that large LRs do not necessarily favour H_P , the frequentist approach alone is inadequate as a measure of the strength of forensic evidence.

Put another way, the frequentist approach to evidential value considers only the rarity of the class characteristic, which is associated with only one (of two) hypotheses in the adversarial legal system. It does not consider the rarity of the other (competing) hypothesis. In our card example, when determining the probability that an unknown card is an ace of spades, a frequentist would only consider the rarity of a spade or an ace, and not the probability that the card is NOT the ace of spades. We showed that while this leads to an LR greater than one, the posterior odds remain less than one (and thus favour the hypothesis that the card is NOT the ace of spades). This neglect was evident in the case of Sally Clark [16] who was wrongfully convicted (and later acquitted) of murdering her two sons. The prosecution case relied on a significantly flawed frequentist interpretation presented by paediatrician Professor Sir Roy Meadow who calculated the rarity of sudden infant death syndrome (SIDS). It was heavily criticised by the Royal Statistical Society [17] on the grounds that, while double SIDS is very rare, double infant murder is likely to be rarer still, so the probability of Clark's innocence was actually quite high:

“Two deaths by SIDS or two murders are each quite unlikely, but one has apparently happened in this case. What matters is the relative likelihood of the deaths under each explanation, not just how unlikely they are under one explanation.”

Further, Professor of Mathematics Ray Hill opined [18]:

“When a cot death mother is accused of murder, the prosecution sometimes employs a tactic such as the following. If the parents are affluent, in a stable relationship and non-smoking, the prosecution will claim that the chances of the death being natural are greatly reduced, and by implication that the chances of the death being homicide are greatly increased. But this implication is totally false, because the very same factors which make a family low risk for cot death also make it low risk for murder.”

5. Conclusions

A frequentist interpretation of the strength of forensic evidence will always underestimate support for the defence hypothesis that a random thing (person, tool, weapon, car, *etc*) could have been the source of forensic evidence. This is particularly the case when only one discrete class characteristic of the evidence is considered (*eg.* colour of a paint chip, type of footwear impression, presence or absence of a single chemical component). It is only required to demonstrate the occurrence of one other thing with this property in a reference collection (or, indeed, perhaps, in the world) for the Bayesian posterior odds ratio to favour the defence hypothesis.

The only way to circumvent this problem is to analyse more than one *independent* class characteristic and to combine the frequencies in the calculation of the LR such that the $LR > N - 1$ (*ie.* the LR is

greater than the number of things observed, all things being equally likely to be the source of the evidence). This has been successfully achieved in forensic genetics by the use of the so-called product rule where the frequencies of genotypes at loci sufficiently separated in the genome as to be independently inherited can be combined as independent events, thus producing LR_s far in excess of the number of people in the world. This approach may be appropriate for other fields. For example, if it can be demonstrated that fingerprint patterns are independent of ridge characteristics, then the frequencies of patterns and ridge characteristics in reference populations could be combined to similarly produce LR_s in excess of the number of people possibly able to have contributed the fingerprint (conservatively, the number of people in the world).

Finally, the frequent inability to estimate the prior odds ratio warrants one or both of the following to be applied to Bayesian evidence interpretation:

- (a) The worst case scenario for the prosecution, ceding all reasonable uncertainty to the defence, should be used to estimate the prior odds ratio. This is consistent with the common judicial prerogative to minimise wrongful convictions (at the expense of wrongful acquittals)
- (b) Prosecution on the basis of a single piece of evidence using a single, discrete class characteristic is fraught with danger and multiple, corroborating evidence items should always be the goal of a prosecution.

Conflicts of interest

None

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ref.: Ms. No. FSI-D-18-00227

Bayesian interpretation of discrete class characteristics Forensic Science International

Author contribution statement

Dennis McNevin is the sole author of this work and contributed all of the:

- Conceptualization
- Methodology
- Validation
- Formal Analysis
- Investigation
- Resources
- Data Curation
- Writing – Original Draft
- Writing – Review & Editing
- Visualization
- Project Administration

References

- [1] J.S. Buckleton, J.-A. Bright, D. Taylor, *Forensic DNA evidence interpretation*, CRC press, 2016.
- [2] I.W. Evett, B.S. Weir, *Interpreting DNA evidence: Statistical genetics for forensic scientists*, Sinauer Associates, Inc., 1998.
- [3] L. Foreman, C. Champod, I. Evett, J. Lambert, S. Pope, *Interpreting DNA evidence: A review*, *International Statistical Review / Revue Internationale de Statistique*, 71 (2003) 473-495.
- [4] I.W. Evett, *A Bayesian approach to the problem of interpreting glass evidence in forensic science casework*, *Journal of the Forensic Science Society*, 26 (1986) 3-18.
- [5] C.G. Aitken, F. Taroni, *Statistics and the evaluation of evidence for forensic scientists*, Wiley Online Library, Chichester, 2004.
- [6] M.O. Finkelstein, W.B. Fairley, *A Bayesian approach to identification evidence*, *Harvard Law Review*, 83 (1970) 489-517.
- [7] T. Bayes, R. Price, *An essay towards solving a problem in the doctrine of chances*. By the late Rev. Mr. Bayes, F.R.S Communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S., *Philosophical Transactions (1683-1775)*, 53 (1763) 370-418.
- [8] F. Taroni, C.G.G. Aitken, P. Garbolino, *De Finetti's subjectivism, the assessment of probabilities and the evaluation of evidence: A commentary for forensic scientists*, *Science & Justice*, 41 (2001) 145-150.
- [9] R.E. Kass, A.E. Raftery, *Bayes factors*, *Journal of the American Statistical Association*, 90 (1995) 773-795.
- [10] H. Jeffreys, *Some tests of significance, treated by the theory of probability*, in: *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, 1935, pp. 203-222.

- [11] H. Jeffreys, *The theory of probability*, OUP Oxford, 1998.
- [12] W.C. Thompson, E.L. Schumann, Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy, *Law and Human Behavior*, 11 (1987) 167.
- [13] D.J. Balding, P. Donnelly, The prosecutors fallacy and DNA evidence, *Criminal Law Review*, (1994) 711-721.
- [14] L.A. Foreman, A.F. Smith, I.W. Evett, Bayesian analysis of DNA profiling data in forensic identification applications, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 160 (1997) 429-459.
- [15] R. Meester, M. Sjerps, Why the effect of prior odds should accompany the likelihood ratio when reporting DNA evidence, *Law, Probability and Risk*, 3 (2004) 51-62.
- [16] *R v. Clark*, in, *EWCA Crim 54*, 2000.
- [17] Royal Statistical Society, News release: Royal Statistical Society concerned by issues raised in Sally Clark case, in, *Royal Statistical Society*, 2001.
- [18] R. Hill, Multiple sudden infant deaths – coincidence or beyond coincidence?, *Pediat Perinat Edipem*, 18 (2004) 320-326.

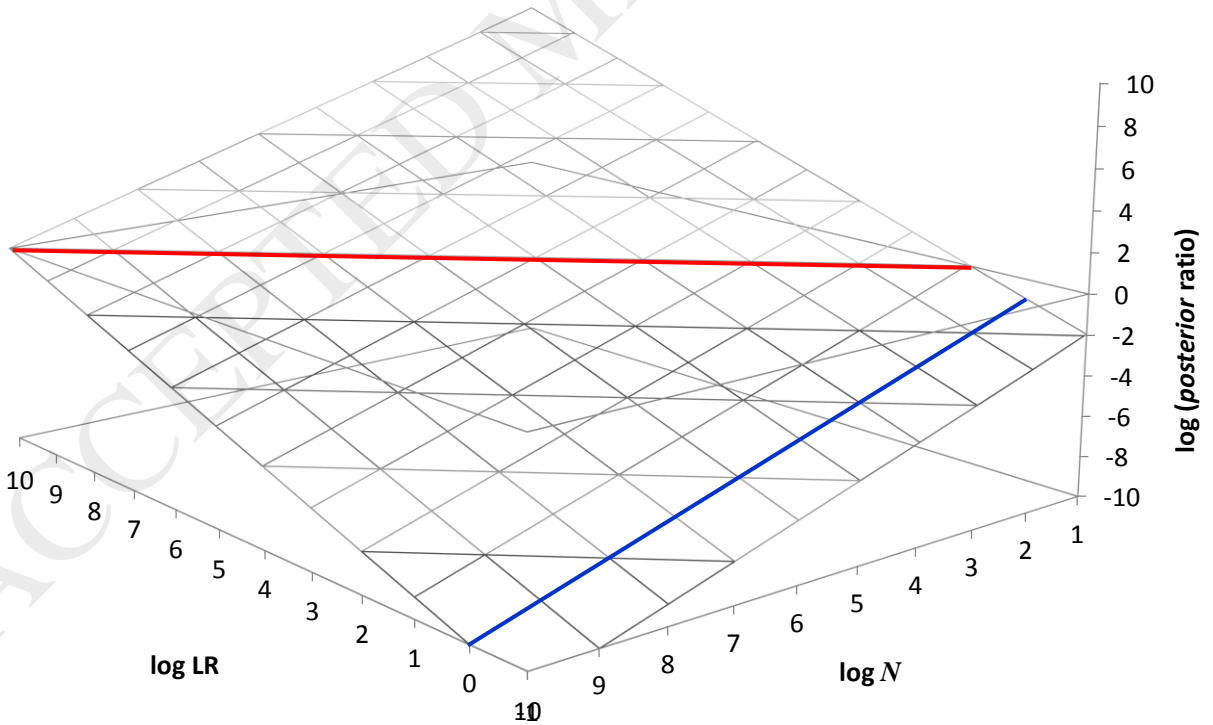
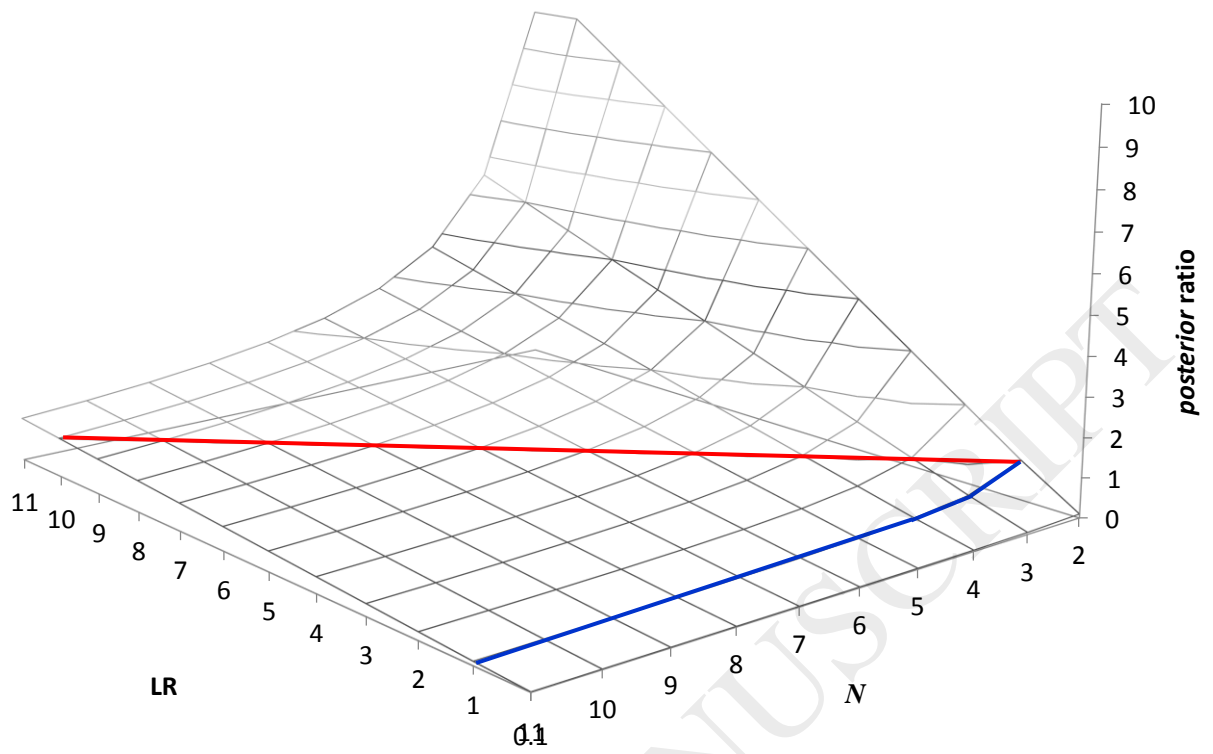


Figure 1 Posterior ratio as a function of both the likelihood ratio (LR) and the number of things that could possibly be the source of the evidence (N), represented with linear (top) and logarithmic (bottom) scales. The red line represents a situation where neither the prosecution or defence hypotheses are favoured given the evidence (posterior ratio = 1 or log posterior ratio = 0) and the blue line represents a situation where the evidence is no more likely under either the prosecution or defence hypotheses ($LR = 1$ or $\log LR = 0$). These lines diverge as N increases.

Table 1 Posterior ratio of probabilities as a function of the number of things (n) that have the same discrete class characteristic

n	$\frac{P(H_p/E)}{P(H_D/E)}$
1	∞
2	1
3	1/2
4	1/3
5	1/4
.	.
.	.
.	.
n	$1/n$