Queen Mary
University of London

The Island Problem

Bayesian Reasoning in the Legal System

Stephen Dewitt
s.dewitt@qmul.ac.uk

1. The Island Problem

Imagine that you live on this Island.
100 other people live there with you in harmony. Until one day ...
A man is found dead. Murdered.

This is bad news for you, as you are the local law enforcer for the Island. You examine the crime scene and find footprints leading away which could only have been left by the murderer. You measure the footprints and find that they are size 12. Checking an anonymous database from the last 100 years encompassing yours and other neighbouring islands, you find out that on average only 1 in 20 islanders tend to have that size shoe.

Somehow (as it always does) word gets around the island about the footprints and as you walk back to the town you see a group of islanders pushing a man in front of them coming towards you. When they reach you they push the man onto the ground in front of you and tell you that he has size 12 shoes and that he must be guilty.

Given that this is the only evidence against the man, what do you think is the chance that he is guilty of the crime? Make your choice below.

Take a Sticker and Make Your Choice

1 in 100 | 1 in 20 | 1 in 6 | 1 in 5

2. Could the Island Problem be Valuable in Law?

Misuse of statistics in courts is leading to miscarriages of justice

Due to increasing computational power forensic teams are increasingly presenting evidence in court cases in a statistical format. While legal professionals typically receive no statistical training they regularly use this evidence in their arguments without professional aid. These combining factors have led to a situation in which statistical fallacies in the UK legal system are frequent, increasing and have led to notorious and disastrous miscarriages of justice (e.g. Sally Clark; see Forrest, 2003).

The statistical community is in agreement that the correct application of 'Bayesian' statistics guided by an expert would eradicate these fallacies, however the presentation of the results of these calculations has proven highly problematic due to their complexity and the low average level of statistical expertise of both legal professionals and jurors (Fenton, Neil & Hsu, 2011).

A recent, highly-criticized attempt to show the jury the fundamental calculations and have them complete them by hand proved disastrous (R v T, 2009; see Aitken & Berger, 2011).

Could a novel approach provide a solution?

Fenton, Neil & Hsu (2013) argued that the combined issues of statistical complexity and lack of expertise in both legal professionals and jurors make any similar future attempts to RvT (2009) fruitless. They instead proposed a new system in which the mathematical calculations involved in a court case would simply be trusted in the same way that one would trust a calculator.

Importantly, both the quantification of evidence itself as well as the decision of guilt or innocence would still be decided in the traditional manner. Only the calculations themselves would be trusted, and the authors argued that these are mathematical fact and should be regarded as such.

In order to promote this required trust as well as a degree of understanding the authors proposed that **the Island Problem**, a simple Bayesian problem with a 'legal framing' could serve as an 'introduction' to Bayesian thinking.

The present research programme is therefore focused on finding the optimum presentation format for the Island Problem in order to promote trust and understanding in more complex Bayesian calculations presented in court cases.

3. How well do People do on the Island Problem?

The Correct Answer

'1 in 6' is strictly the correct answer to the Island problem. However '1 in 5' has historically been accepted as correct in previous studies. It has been accepted because the process required to get to '1 in 5' has generally been of greater cognitive interest than the extra step then required to get to '1 in 6'.

The reason that '1 in 6' is correct is because we are told that the resident has size 12 shoes and so it is not correct to apply the 1 in 20 probability figure to him (the chance of him having size 12 shoe is 1 in 1). Therefore we apply the '1 in 20' figure only to the remaining 100 islanders (which equals 5) and then add the defendant on (which equals 6) to get the most likely total number of islanders with that shoe size.

The defendant is 1 of these 6 people who have size 12 shoes and there is no other evidence against him so '1 in 6' is the best estimate of the chance that he committed the crime.

A Preliminary Experiment

- **We presented people with a version of the Island problem** with no instructions to get a baseline of how well the public do on the problem. We also asked them to explain their thought process in text format.
- **No one gave the answer of 1 in 6**, which is strictly correct.
- **57% gave the answer of 1 in 5**, which has historically also been accepted as correct.
- **28% of people choose the incorrect answer of '1 in 20'.** This is a well-known finding, and previous work which has largely not examined thought processes has assumed that all people who give this answer commit the same mistake: they have assumed that all such people 'neglect the base rate' (the '100' population figure is the base rate), focusing only on the '1 in 20' figure.
- **However, we found three distinct thought processes which led participants to make the '1 in 20' mistake** from the text responses. Explanations of these three processes are given to the right, as well as a selection of quotes from the text responses which epitomize the category.

Mis Readers — *1 in 20* ☒

These individuals appeared to simply mis-read the '1 in 20 Islanders have size 12' figure as '20 islanders have size 12'. Given this mistake, their subsequent calculation is correct.

"If there are only 20 [islanders] who wear a size 12, then the chance of the suspect being guilty is 1/20"

"Twenty [islanders] wear a size 12, and he is one of them"

Base Rate Neglecters — *1 in 20* ☒

These individuals actually seem to fit the single historically-assumed cause: their text responses indicate that they simply ignored the '100' population figure.

"The odds of [an islander] having a size 12 shoe is 1 in 20. Since the defendant had that size, the same odds apply."

"The suspect had a size 12 shoe and the probability is 1 in 20 so that means that this is the same probability that he did it."

Reversers — *1 in 20* ☒

These individuals are perhaps the most intriguing. They undertook all the necessary steps to get to the answer of '1 in 5', and then, curiously, 'reversed' the calculation to get back to '1 in 20'.

"There were 100 so 5 in 100 have size 12 shoe. Brings it back to 1/20."

"Out of a group of 20, 1 is expected to have size 12. 5 out of 100 reduces to 1 out of 20."

4. Next Steps

1 Greater Understanding

The first aim of the present research programme is to fully understand why people fail to give normative answers to the Island Problem. The text-responses in the preliminary experiment were a good first step towards this aim. An important next step will be to replicate this work while expanding the depth and details of responses possible by participants to gain a greater understanding.

2 Promoting Trust

Based on the data from the first experiment, we will design an 'Island' tutorial presentation and determine whether this produces greater trust and understanding in the Bayesian statistics used in court cases compared to competing formats.

3 Optimisation

The presentation format divined from experiment 2 will be further refined and explored through experimentation to produce the optimal format. Further mixed methods data and eye tracking will be utilized to determine *why* the presentation aids understanding and who it works best for.

References

Aitken, C., & Berger, C. (2011). Expressing evaluative opinions: a position statement. *Science & Justice*
Fenton, N., & Neil, M. (2011). Avoiding Probabilistic Reasoning Fallacies in Legal Practice using Bayesian Networks, (June), 1–44.
Fenton, N., Neil, M., & Hsu, A. (2013). Calculating and understanding the value of any type of match evidence when there are potential testing errors. *Artificial Intelligence and Law*, (September), 1–25.
Forrest, A. (2003). Sally Clark–A lesson for us all. *Science & Justice*.