Using Bayes to prove Obama did not write his own book?

Norman Fenton, 18 July 2011

There have been many questions about the closeness of President Obama's relationship with Weather Underground terrorist Bill Ayers. A whole new angle on the relationship has been raised in Jack Cashill's book Deconstructing Obama. Using information in this book Andre Lofthus has applied Bayes Theorem to conclude that Bill Ayers actually was the ghost writer for Obama's best selling book Dreams from My Father.Loftus's analysis is based on

a) a comparison of *Dreams* with one of Ayers's own books *Fugitive Days*; and

b) a comparison of *Dreams* with a different book *Sucker Punch* based on similar material to that of both *Dreams* and *Fugitive Days*.

Specifically in a) there were 759 similarities, of which 180 were categorized by Cashill as "striking similarities", whereas in b) Cashill claims there were just six definite similarities, with a maximum of sixteen possible or definite similarities. Lofhus's Bayesian analysis leads him to conclude with certainty that the books *Dreams* and *Fugitive Days* cannot have been written by different people and that, since there was never any suggestion that Ayers was not the author of *Fugitive Days* then he must be the author of Dreams.

His Bayesian analysis is oversimplified and contains some unrealistic assumptions, so I have done my own analysis here. My own conclusions are not as definite. The evidence does indeed provide very strong support in favour of the books being written by the same author. However, if you have a strong prior belief that the books were written by different authors (say you are 99.9% sure) then even after observing the evidence of 180 striking similarities, it turns out that (with what I believe are more reasonable assumptions than made by Lofhus) there is still a better than 50% chance that the books were written by different authors.

The analysis

Cashill considers two competing hypotheses:

H1: "Obama is the author of *Dreams*".
H2: "Ayers is the author of *Dreams"*

and the evidence

e: "Discovery of 180 striking similarities between *Dreams* and *Fugitive Days"*.

One minor problem with his analysis is that Lofhus assumes that H1 and H2 are mutually exclusive and exhaustive for the Bayesian calculations but later accepts that there could be a third alternative hypothesis that a third author may have written *Dreams*. Since we are only really interested in whether someone *other* than Obama wrote it, it is simplest to assume just the two hypotheses:

H: "*Dreams* and *Fugitive Days* were written by different authors".

not H: "*Dreams* and *Fugitive Days* were written by the same author".

This enables us to consider, as Lofhus does, the likelihood ratio:

> P(e | H) divided by P(e | not H).

If the likelihood ratio is above one then the evidence provides support for H and if it is below 1 it provides support for not H.

According to Lofhus, P(e | H) is as good as equal to 0.

> ***Comment on this***: Lofhus's justification for this calculation is that, for books with different authors, the number of striking similarities is a *Normal distribution with mean 6 and variance 100*. So, if H were true, then he assumes that the number of striking similarities between these two books would also be drawn from the same Normal distribution. In that case, the probability of actually seeing the evidence e (180 striking similarities) is indeed extremely close to zero as claimed. However, it is more likely that the real distribution of striking similarities between books with different authors is not the very 'narrow' Normal distribution assumed. It is likely to be a distribution with a small, but fairly long tail. In that case P(e | H) will be small, but non-zero.

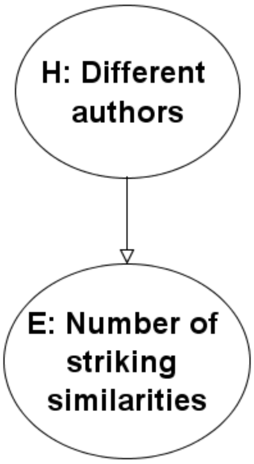
According to Cashill P(e | not H) is equal to 1.

> ***Comment on this***: This is NOT correct. To calculate P(e | not H) we have to have a prior distribution for the number of striking similarities between two books written by the same author. Lofhus does not provide any suggestion of what this distribution might be. Even if we did know what the prior distribution was, P(e | not H) is certainly not equal to 1. For example, even if we suppose that the prior distribution is a Normal distribution with mean 180 and a small variance, say 100, then the probability that exactly 180 striking similarities is found is actually about 0.04. If we assume the prior distribution has a greater variance, say a Normal with mean 180 but a variance of 2000 (much more realistic) then the probability of seeing exactly 180 similarities is about 0.011.

However, even if we take account of Lofhus's error on P(e | not H) his overall conclusion, (using his assumptions about the Normal distribution for books by different authors) is actually sound. That is because, although P(e | not H) is quite low, say 0.011 with the second assumption I have used above, it is nowhere near as low as P(e | H). This means that the evidence massively swings the odds away from H in favour of not H. Unless we had an exceptionally strong prior belief (essentially we would need to be certain) that the books were written by different authors, the evidence would leads us to conclude that they were written by the same author.

So are Lofhus's assumptions reasonable and does it really matter? A proper Bayesian analysis requires us to look at a broader range of assumptions than Lofhus has made.

To do this we build a simple two-node Bayesian net model (as shown here using AgenaRisk):



The node H is a Boolean node, but the node E (number of striking similarities) is numeric – instead of assuming E is Boolean (either 180 or not) we have to consider it as a *distribution*.

The prior information we need to specify to run the model is:

P(E | H) (what is the distribution of E for books written by different authors)

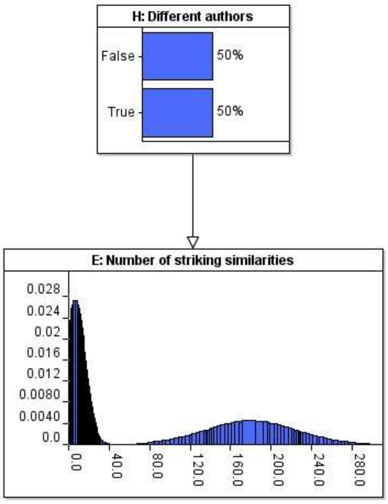
and

P(E | not H) (what is the distribution of E for books written by the same author)

As discussed above Lofhus assumes that P(E | H) is a Normal distribution with mean 6 and variance 10, whereas he never provides a distribution P(E | not H).

In AgenaRisk we can easily run the model and compute the results with a wide range of different assumptions.

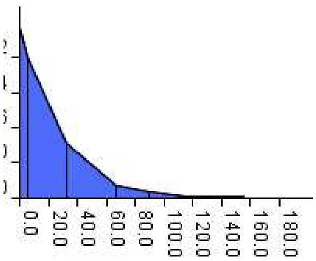
For example, using distributions I discussed above, namely Normal(6,100) for E|H and Normal(180, 2000) for E | not H, the model (before any evidence is entered) is as follows:



This assumes a prior of 50:50 between H and not H.

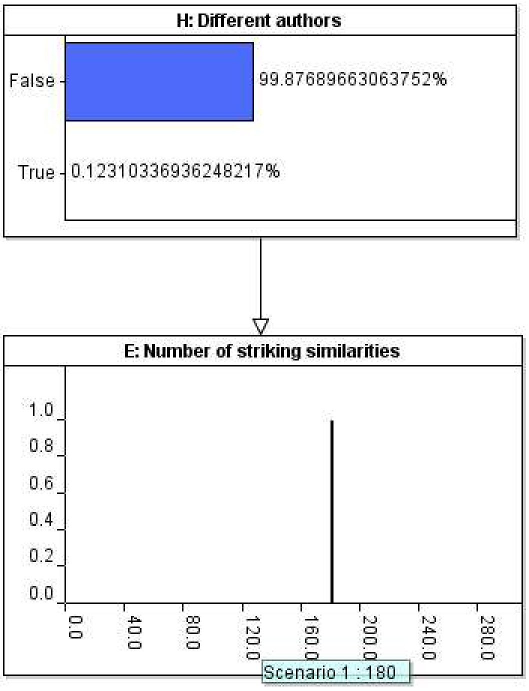
Because of the inevitable approximations used for exceptionally low numbers, the probability that E can be as high as 180 when H is true is calculated to be 0 in this case, so if we enter the evidence E=180 and run the model we get a posterior probability that H is 1. Moreover, even if we make the prior for H being true 'almost 1' (say 0.9999999999999) then the model still returns a posterior of H being false as 1 when we enter the evidence E=180. So, with these assumptions the model does indeed support Lofhus's conclusions.

But if we change the assumptions we get very different results. Suppose, for example, we assume that P(E | H) is a distribution that has a small, but long tail (this seems more reasonable) such as this:

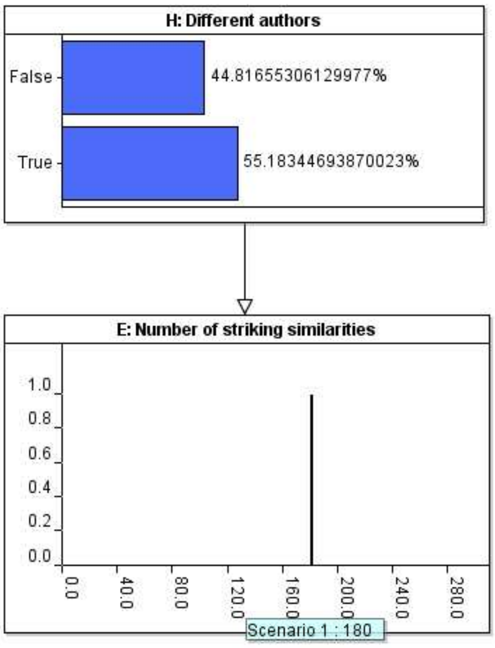


(this is a beta(1, 80, 0, 200) distribution)

Then when we run the model entering the evidence E=180 we get the following result:



i.e. assuming the prior of 50:50 between H being true and false, there is clearly a non-zero probability (albeit less than 1%) that the authors are different. But if, as seems reasonable, we start with quite a strong prior belief that the books have different authors (say we assume P(H) = 0.999), then even after observing the evidence E=180 it is still more likely that the authors are different:



Caveats

- I have not taken account of a number of additional subjective factors that support the books both being written by Bill Ayers (e.g. the number of similarities not considered 'striking').
- What is really needed is more extensive data to determine the real distributions for E|H and E| not H. In practice I suspect there will be more variance in these distributions than I have assumed above (there will certainly be much more variance than assumed by Lofhus). In that case the single piece of evidence E=180 will not necessarily provide overwhelming proof that the books were written by the same author, even if we assume a prior of 50:50 between H and not H. For example, the following table shows the posterior probability of H being true after observing E=180 for similar, but increasingly 'flat' distributions:

E | H	E | not H	Prob H after observing E=180
beta(1, 80, 0, 200)	Normal(180, 2000)	0.12%
beta(1, 50, 0, 200)	Normal(180, 5000)	2.8%
beta(1, 30, 0, 200)	Normal(180, 10000)	16.4%
beta(1, 20, 0, 200)	Normal(180, 20000)	34.6%

The simple Bayesian net model is here (you need to download the free trial version of AgenaRisk to run it).

Return to Probability and Risk page