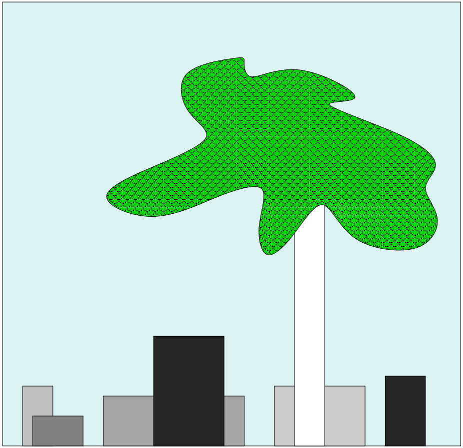
Occam's Razor

Thomas Bayes

April 28, 2018

Question

How many boxes are behind the tree? One or two?



Occam's Razor

- *Accept the simplest explanation that fits the data.*
- *Plurality is not to be posited without necessity.*
- *Less is more.*

The intuition maybe is 'well, it would be a remarkable **coincidence** for the two boxes to be just the same height and colour as each other.'

Bayesian derivation of Occam's Razor

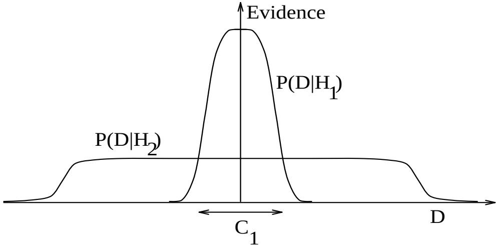
We have two hypothesis $\mathcal{H}_1$ and $\mathcal{H}_2$. We are interested in the posterior of $\mathcal{H}_1$ and $\mathcal{H}_2$ in light of data D.

$$\frac{P(\mathcal{H}_1|D)}{P(\mathcal{H}_2|D)} = \frac{P(\mathcal{H}_1|D)}{P(\mathcal{H}_2|D)}\frac{P(D|\mathcal{H}_1)}{P(D|\mathcal{H}_2)}$$

The first part is our preference of the one-box or two-box hypothesis. In this case, it seems quite natural to say we have no prefrences over the number of boxes(Note that Occam's Razor principle does not apply here, because one-box hypothesis does not necessarily mean simpler.)

One box or two box?

Our two-box hypothesis will lay wider distribution on the space, thus the *particular data* we are observing will share lower probablity.



Number example

In this example (extracted from online), we have two hypothesis, $\mathcal{H}_1$: the sequence is an arithmetic progression, $\mathcal{H}_2$: the sequence is generated from a fourth-order polynomial.

One possible reason for our preferences over $\mathcal{H}_1$ is maybe, in our mind, we think a sequence is more likely to be generated from arithmetic progressio, so $P(\mathcal{H}_1) > P(\mathcal{H}_2)$.

And another reason is that it is less likely for a fourth-order polynomial to somehow accidently generate four integer numbers!

References

Chapter 28 of *Information Theory, Inference, and Learning Algorithms* by **David J.C. MacKay**