SETS, RELATIONS AND PROBABILITY LECTURE 6

1. We saw that the definition of conditional probability was this: Pr $(A \mid B)$ = Pr $(A \cap B)$ / Pr (B). We used this to calculate conditional probabilities, but we can also use it the other way around: given Pr $(A \mid B)$ and Pr (B), we can calculate Pr $(A \cap B)$ by means of the formula:

   Pr $(A \cap B)$ = Pr $(A \mid B) \times$ Pr (B)

2. For instance, suppose we know that: exactly 50% of the population is male and that 30% of males have blue eyes. Then what is the probability that a randomly drawn person is male *and* blue-eyed? Letting M say that the person is male and B say that they are blue-eyed, we are asking after Pr $(B \cap M)$:

   (i) Pr (M) = 0.5
   (ii) Pr $(B \mid M)$ = 0.3
   (iii) Pr $(B \cap M)$ / Pr (M) = Pr $(B \mid M)$
   (iv) Pr $(B \cap M)$ / 0.5 = 0.3
   (v) Pr $(B \cap M)$ = 0.15

3. Suppose we take probability to measure one's confidence in a proposition. Then the **Bayesian** claim is that after you learn some new bit of evidence E, your new confidence in any hypothesis H is your old Pr $(H \mid E)$. For instance, in the preceding example, your confidence that a randomly drawn person has blue eyes, once you learn they are male, should be 30%.

4. One crucial application of the Bayesian thesis is to cases where one knows Pr (H), Pr (E) and Pr $(E \mid H)$ and then learns that E is true. How confident should one *then* be that H is true? The answer is Pr $(H \mid E)$, and we can calculate this using the following rule:

   Bayes's Theorem (first version): Pr $(H \mid E)$ = Pr $(E \mid H) \times$ Pr (H) / Pr (E)

   What this means is e.g. that if some hypothesis either predicts, or makes highly likely, some observation that was very unlikely in advance, then actually making that observation is very strong evidence for the hypothesis.

5. For instance, suppose that I hold some bizarre conspiracy theory H according to which the end of the world will occur on 5 March 2015, and one consequence of which is (E) that tomorrow there will be earthquakes all over the world. We might expect Pr (E) and Pr (H) both to be very small today, and Pr $(E \mid H)$ = 1 (make sure you are clear why). But if we do observe E tomorrow, then Bayes's Theorem implies that Pr (H) should now rise enormously, in fact that it should increase by a factor 1/Pr (E).

6. Again, consider the Monty Hall puzzle. You (A) or one of two other prisoners (B, C) will be shot tomorrow. You ask the guard to tell you which of B and C

will *not* be shot (or to mention one at random if neither will). The guard tells you truthfully that B will not be shot. How confident should you now be that *you* won't be shot? Letting H be the event that you *will* be shot, and E the event that the guard mentioned B, we have Pr (E|H) = 0.5, Pr (H) = 1/3 and Pr (E) = 0.5. The formula therefore gives Pr (H|E) = 1/3: the information should make you twice as confident that C will be shot as that you will.

7. We can extend this use of Bayes's Theorem to cases where you don't directly know Pr (E) but you do know both Pr (E|H) and Pr (E|¬H). In that case, we use the identity of probability theory (which you should know but need not be able to prove): for any E and H:

   $$\Pr(E) = \Pr(E|H)\Pr(H) + \Pr(E|\neg H)\Pr(\neg H)$$

   Substituting this into the first version of Bayes's Theorem we get:

   Bayes's Theorem (second version):

   $$\Pr(H|E) = \frac{\Pr(E|H)\Pr(H)}{\Pr(E|H)\Pr(H) + \Pr(E|\neg H)\Pr(\neg H)}$$

   This formula is useful for answering questions like this one: if we have a detector of given reliability, how confident should we be in the proposition that it detects for given that it is positive?

8. Probably the most famous application is Hume's treatment of miracles. Suppose (E) that a witness (say, an apostle) says that a man walked on water. Suppose we start out with Pr (H) = 1 in a million, Pr (E|H) = 90% and Pr (E|¬H) = 1%. (So the witness is highly reliable.) Then how likely is it that the man did walk on water? Writing 1M for a million, the formula gives us:

   $$\frac{0.9/1M}{0.9/1M + (0.01 \times 999,999)/1M} \approx 1/10000$$

9. Notice the following about the formula in 7. Pr (H|E) will never exceed 0.5 unless Pr (E|H) Pr (H) exceeds Pr (E|¬H) Pr (¬H). But this means that Pr (E ∩ H) > Pr (E ∩ ¬H), hence only if false testimony is less likely than the reported miracle itself. That is why Hume wrote: '[n]o testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish.'