SETS, RELATIONS AND PROBABILITY LECTURE 6

1. We can extend 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), here writing ¬H for 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 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 very 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.

2. 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 taken to be highly reliable.) Then how likely is it that the man did walk on water? Writing 1M for a million, the formula gives us:

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

3. We now consider probabilities of (sets of) outcomes of *repeated* tests of devices or arrangements with random results. Suppose that we have a reference set V whose members are the possible outcomes of some trial, and we wish to repeat the experiment once. Then if the possible outcomes as the same in both cases, we may consider the repeated test as *one* trial with each *ordered pair* of V corresponding to an element of the new reference set, written $V \times V$ or $V^2 = \{(x, y): x \in V \wedge y \in V\}$ (recall lecture 2 no. 3).

4. First we consider the case where the outcome of each individual trial is *probabilistically independent* of the outcome of any other. Thus suppose that you toss a coin twice, and suppose that the tosses are probabilistically independent, so that your confidence in the result of any toss is completely unaffected by your knowledge of the outcome of any distinct toss. (This is in fact something of a fiction.) Then the outcomes of each toss are elements of V = {H, T}. And the outcome of both is an element of V^2 = {(H, H), (H, T), (T, H), (T, T)}. Hence there are 4 possible

outcomes: ignoring the parentheses we can write these as: HH, HT, TH, TT. Each outcomes gets probability ¼. What is the probability that you get heads at least once? It is P ({HH, HT, TH})= ¼ + ¼ + ¼ = ¾.

5. Now consider a case of repeated trials where the outcome of each is *probabilistically dependent* on that of the others e.g. random draws from a pack of cards without replacement. Consider the case where two draws are made. Then our initial reference set V has 52 members representing one draw. The new reference set is not V^2 because you can't remove the same card twice. Instead it is $V^2 - X$, where $X = \{(x, x): x \in V\}$. That set has 51×52 members each of which (i.e. each of whose one-element subsets) has equal probability. Then what is the probability that one of them is an ace? Consider the following sets of events:

   A: The first is an ace and the second is not: 4×48 elements
   B: The second is an ace and the first is not: 4×48 elements
   C: They are both aces: 4×3 elements

   Then what we want is $\Pr(A \cup B \cup C)$. We know that $\Pr(A \cup B \cup C) = \Pr(A) + \Pr(B) + \Pr(C)$ by axiom (iii), since any two of A, B, C have an empty intersection. So the answer is:

   $$\Pr(A \cup B \cup C) = (4 \times 48 + 4 \times 48 + 4 \times 3) / (51 \times 52)$$
   $$= 99 / (13 \times 51) = 33 / (13 \times 17) = 33 / 221$$

6. We can also calculate *conditional* probabilities of results of repeated trials. Suppose that we are tossing a fair coin three times. Then there will be 8 outcomes of equal probability viz HHH, HHT, HTH, HTT, THH, THT, TTH, TTT. What is the probability that they are all heads given that one of them is heads? The quantity we want is:

   Pr ({HHH}|{HHH, HHT, HTH, HTT, THH, THT, TTH})
   = Pr ({HHH} ∩ {HHH, HHT, HTH, HTT, THH, THT, TTH}) / (7/8)
   = (1/8) / (7/8) = 1/7

7. Now suppose we want to calculate the probability that they are all heads given that the *first* toss lands heads. In this case we want:

   Pr ({HHH}|{HHH, HHT, HTH, HTT}) = (1/8) / (4/8) = 1/4

   Notice that this answer is about twice as large as the last one. But how can this be? The results of the tosses are completely (i.e. both causally and probabilistically) independent, so if somebody notices that one of the three tosses landed heads, why should the additional information that the tossed he noticed happened to be the *first* one. Why should that make any difference to what you think happened on the other two?