Example problems using Bayes' Theorem:

Bayes' Theorem says that P(H|O) = P(O|H) * P(H)/P(O). Since we know that P(O) = P(H&O) + P(H&~O) = P(O|H)*P(H) + P(O|~H)*P(~H) we sometimes just expand the denominator and write:

$$P(H|O) = \frac{P(O|H) * P(H)}{P(O|H)*P(H) + P(O|\sim H)*P(\sim H)}$$

Bayes' Theorem is often used for 'inverse inference' where we have a good model of the 'direct' probabilities – that is, the likelihoods which are of the form P(O|H).

For example, in medical diagnosis and testing, we might know the error rates of a testing device. Lets say that 'D' means that the patient has the disease while '+' means that the test result was positive. The 'false positive rate' is P(+|~D) - the probability of getting a positive result when the patient does not have the disease. The false negative rate is P(-|D). Note that these have no necessary connection to each other – for example, they do not have to be the same nor do they have to sum to one. A good medical test might have a false positive rate of .01 and a false negative rate .05 for example. Though because of the way medical testing works, it is usually the case that it would be possible to design a more sensitive test that misses less (has a lower false negative rate) but this would cause the test to have a much higher false positive rate. When sensitivity and reliability trade off in this way, there is no apriori way to determine which is the better kind of test. It simply depends on the costs and benefits of correct vs. incorrect results in the particular case.

Example: Imagine that a given patient has a 25% chance of having a certain disease. You give the patient a test which comes up with a positive result. The test has a 1% false positive rate and a 2% false negative rate. What is the probability that the patient actually has the disease?

Answer: Since we do not know P(+) directly, we have to use the expanded form of Bayes' Theorem

$$P(D|+) = \frac{P(+|D) * P(D)}{P(+|D)*P(D) + P(+|\sim D)*P(\sim D)}$$

Here we know that P(~D) = 1-P(D) and (+|D) = 1-P(-|D). Thus we have:

$$\frac{P(+|D) * P(D)}{P(+|D)*P(D) + P(+|\sim D)*P(\sim D)} = \frac{.98 * .25}{(.98 * .25) + (.01 * .75)} = .97$$

So in this case, you can be reasonably sure (but not certain!) that the patient does have the disease.

However, in apparently similar problems, this would not be the case. If the base-rate (=prior probability) that the patient has the disease is very, very low relative to the error rate, then it can actually be more probable that the positive test result is a false positive rather than a true positive. For example, if roughly 1/1000 randomly selected people would have the disease and 1/100 people who do not have the disease would test positive, then if you test 1000 randomly selected people, you can expect to get about 11 positive results only 1 of whom actually has the disease. Thus the probability that a given person who tests positive actually has the disease is roughly 1/11 (which is way higher than 1/1000 – so the positive test result is definitely evidence – but just very far from conclusive).

Example: Imagine that you have three urns that you cannot see into. Urn1 is 90% green balls and 10% red. Urn2 is 50% green and 50% blue. Urn3 is 20% green, 40% red, and 40% blue. You can't tell which urn is which. You randomly select an urn and then randomly select a ball from it. The ball you drew is green. What is the probability that it came from urn1?

Answer: Here if there were only two urns you could treat one of them as 'H' and the other as '~H' for purposes of uses Bayes' Theorem. But you can't do that here since there are three possibilities.

Bayes Theorem is still true and still says:

$$P(Urn1|G) = \frac{P(G|Urn1) * P(Urn1)}{P(G)}$$

But the problem is that there are now three different ways to get a green ball (Urn1, Urn2, Urn3) and you need to take all of them into account when calculated the prior odds of getting a green ball. Here we just use the more general form of Bayes Theorem which says:

$$P(H|O) = \frac{P(O|H) * P(H)}{\Sigma_n \ P(O|H_n)*P(H_n)}$$

where 'Σ_n' is the summation sign (plug in 1 for n then 2, then 3, etc. for each of the possible ns and then sum all of the possible results). Note that if there uncountably many possible values for n you use an integral.

In this urn case, Bayes theorem says:

$$P(Urn1|G) = \frac{P(G|Urn1) * P(Urn1)}{P(G|Urn1) * P(Urn1) + P(G|Urn2) * P(Urn2) + P(G|Urn3) * P(Urn3)]}$$

Since we randomly selected the urn in the first place, each of P(Urn1), P(Urn2), and P(Urn3) = 1/3. Thus we have:

$$P(Urn1|G) = \frac{.90 * 1/3}{.90 * 1/3 + .50 * 1/3 + .20*1/3} = .5625$$

A parallel calculation shows that P(Urn2|G) = .3125 and P(Urn3|G) = .125. It might be easier to understand why this is correct if we imagine that each urn has exactly 100 balls and thus there are a total of 160 green balls in all three urns and 90 of them are in Urn1 and 90/160 = .5625. But note that these probabilities depend only on frequencies and not absolute numbers of balls at all. If urn1 had 10,000 balls and 9,000 of then were green and urn3 had only 2 green balls out of 10, then frequencies would be the same and thus the final posterior probabilities would also be the same.

It is worth your time to check your understanding to make sure that you see how to get these results:

If you drew a red ball, the probability that it came from Urn1 is .2, Urn2 is 0, and Urn3 is .8. If you drew a blue ball, Urn1 is 0 while Urn2 is 5/9 and Urn3 is 4/9.