Bayes' Theorem

Given two dependent events *A* and *B*, the previous formulas for conditional probability allow one to find *P*(*A* and *B*) or *P*(*B*|*A*). Related to these formulas is a rule developed by the English Presbyterian minister Thomas Bayes (1702 – 61). The rule is known as **Bayes' theorem**.

It is possible, given the outcome of the second event in a sequence of two events, to determine the probability of various possibilities for the first event. In the previous "boxes and balls" example, there were two boxes, each containing red balls and blue balls. A box was selected and a ball was drawn. The example asked for the probability that the ball selected was red. Now a different question can be asked: If the ball is red, what is the probability it came from the box 1? IN this case, the outcome is known, a red ball was selected, and one is asked to find the probability that it is a result of a previous event, that it came from box 1. Bayes' theorem can enable one to compute this probability and can be explained by using tree diagrams.

Bayes' Theorem (cont'd)

The example is repeated here: Box 1 contains 2 red balls and 1 blue ball. Box 2 contains 3 blue balls and 1 red ball. A coin is tossed. If it falls heads up, box 1 is selected and a ball is drawn. If it falls tails up, box 2 is selected and a ball is drawn. Find the probability of selecting a red ball.

To answer the question "if the ball selected is red, what is the probability that it came from box 1?" the two formulas

$$P(B \mid A) = \frac{P(AB)}{P(A)}$$

$$P(AB) = P(B \mid A)P(A)$$

can be used jointly. Finding the probability that box 1 was selected given that the ball selected was red can be written symbolically as $P(B_1|R)$.

Bayes' Theorem (cont'd)

Because

$$P(B_1 \mid R) = \frac{P(B_1 R)}{P(R)} = \frac{P(B_1)P(R \mid B_1)}{P(R)}$$

And *P*(*R*) can be found by using the addition rule:

$$P(R) = P(R \mid (B_1 or B_2)) = P(B_1)P(R \mid B_1) + P(B_2)P(R \mid B_2)$$

Which then gives us

$$P(B_1 \mid R) = \frac{P(B_1 R)}{P(R)} = \frac{P(B_1)P(R \mid B_1)}{P(B_1)P(R \mid B_1) + P(B_2)P(R \mid B_2)}$$

This is a simlified version of Bayes' theorem.

Prior probability or marginal probability of *A*

Likelihood function for *A*

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)}$$

Posterior probability for *A* given *B*

Prior probability of *B*

An Example

Suppose that a test for a particular disease has a very high success rate: if a tested patient has the disease, the test accurately reports this, a 'positive', 99% of the time (or, with probability 0.99), and if a tested patient does not have the disease, the test accurately reports that, a 'negative', 95% of the time (*i.e.* with probability 0.95). Suppose also, however, that only 0.1% of the population have that disease (*i.e.* with probability 0.001). We now have all the information required to use Bayes' theorem to calculate the probability that, given the test was positive, that it is a false positive. Let *A* be the event that the patient has the disease, and *B* be the event that the test returns a positive result. Then, using the second form of Bayes' theorem (above), the probability of a *true* positive is

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)} = \frac{0.001 * 0.99}{0.001 * 0.99 + 0.05 * 0.999} = 0.019$$

An Example

and hence the probability of a false positive is about (1 − 0.019) = 0.981.

Despite the apparent high accuracy of the test, the incidence of the disease is so low (one in a thousand) that the vast majority of patients who test positive (98 in a hundred) do not have the disease. (Nonetheless, this is 20 times the proportion before we knew the outcome of the test! The test is not useless, and re-testing may improve the reliability of the result.) In this case, Bayes' theorem helps show that the accuracy of tests for rare conditions must be very high in order to produce reliable results from a single test, due to the possibility of false positives.