Bayes Theorem

1. Idea

Last time we looked at conditional probabilities. This basically involved calculating the probability of an event given that we know another event has occurred. However, there are situations where knowledge of one event's occurrence does not help (conveys no information) about the occurrence of a second event. This is statistical independence.

2. Recall from last time

$$P(A|B) = \frac{P(A \cap B)}{P(A \cap B) + (A' \cap B)} = \frac{P(A \cap B)}{P(B)}$$

but also

$$P(B|A) = \frac{P(A \cap B)}{P(A \cap B) + (A \cap B')} = \frac{P(A \cap B)}{P(A)}$$

if we rearrange things a bit, we notice:

$$P(A|B)P(B) = P(B|A)P(A) = P(A \cap B)$$

which we could reduce to

$$P(A|B)P(B) = P(B|A)P(A)$$

if we divide both sides by P(B) we get

$P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$ this is also known as Bayes' Theorem

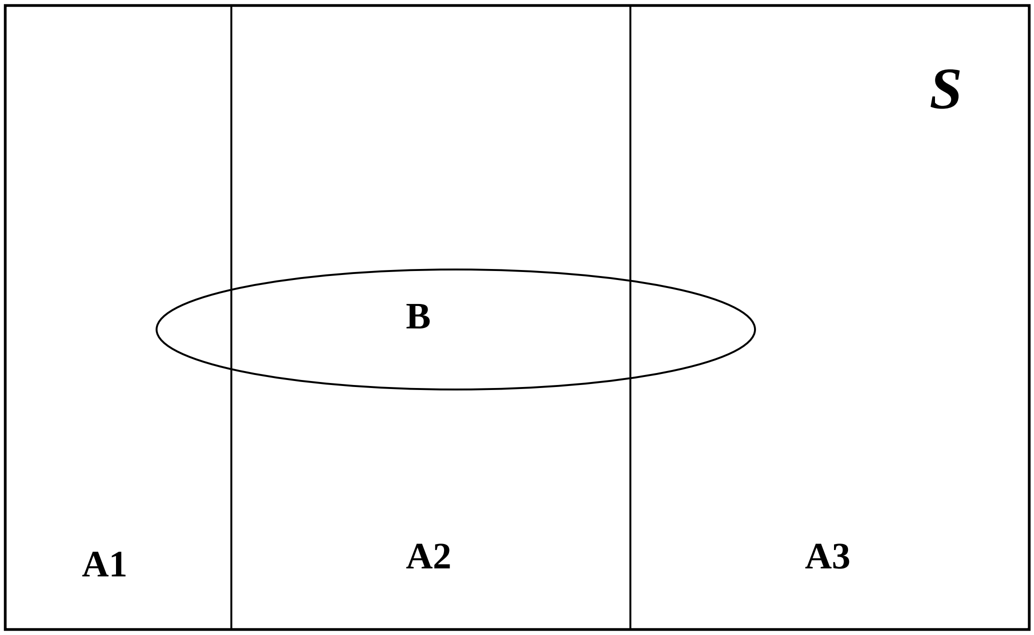
3. A Definition

Each term in Bayes' theorem has a conventional name. The term $P(A)$ is called the *prior probability* of *A*. It is "prior" in the sense that it precedes any information about *B*. Equivalently, $P(A)$ is also called the *marginal probability* of *A*. The term $P(A|B)$ is called the *posterior probability* of *A*, given *B*. It is "posterior" in the sense that it is derived from or entailed by the specified value of *B*. The term $P(B|A)$, for a specific value of *B*, is called the *likelihood function* for *A*. The term $P(B)$ is called the prior or *marginal probability* of *B*.

From the handout (section 2.4, page 17) there are two additional results that will come in handy in understanding Bayes.

Result 7. $P(A \cup B) = 1 - P(A \cup B)' = 1 - P(A' \cap B')$ the probability of A or B occurring is 1 minus the probability of "not A" AND "not B" occuring.

Result 8. Suppose the sample space is divided into $A_1, A_2, \ldots, A_N$ partitions. For any event B then $P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_N)$



Restating result 8: if we added the probabilities of observing all the common outcomes between B and each of the partitions of A, we would get the entire probability of observing B. Result 8 could be rewritten as $P(B) = P(B \cap A) + P(B \cap A')$

This leads to a rewriting of Bayes Theorem

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

Bayes is just a partitioning of the sample space. Your W&W textbook calls it a “tree reversal” and that is a good way to understand Bayes.

4. Examples and Applications

Example 1. In a recent New York Times article, it was reported that light trucks, which include SUV’s, pick-up trucks and minivans, accounted for 40% of all personal vehicles on the road in 2002. Assume the rest are cars. Of every 100,000 car accidents, 20 involve a fatality; of every 100,000 light truck accidents, 25 involve a fatality. If a fatal accident is chosen at random, what is the probability the accident involved a light truck?

Example 2. As accounts manager in your company, you classify 75% of your customers as "good credit" and the rest as "risky credit" depending on their credit rating. Customers in the "risky" category allow their accounts to go overdue 50% of the time on average, whereas those in the "good" category allow their accounts to become overdue only 10% of the time. What percentage of overdue accounts are held by customers in the "risky credit" category?

Things to notice: (1) you have very little information to work with – example 1, you know the fatality rate and the proportion of light trucks, example 2, you know the proportion of good credit customers and the overdue time percentages. (2) A “reverse” tree can help you display the calculations.

Events
A′- Cars
A –Light truck
B –Fatal Accident
B′ - Not a Fatal Accident

Given
P(B|A′) = 20/10000 and P(B|A) = 25/100000

P(A) = 0.4

In addition we know A and A′ are complementary events

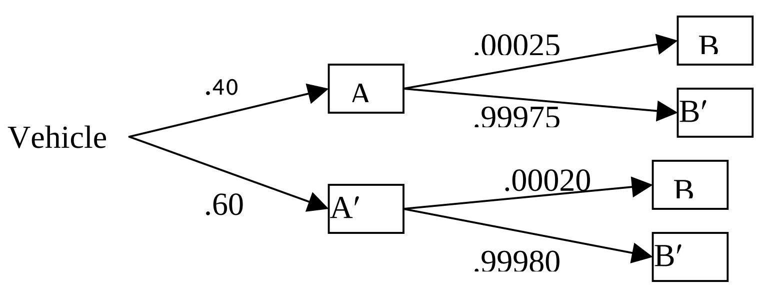
P(A′)=1-P(A)=0.6

Our goal is to compute the conditional probability of a Light truck accident given that it is fatal P(A|B).

Consider *P(A|B)*
Conditional probability of a Light truck involved accident given that it is fatal.
How do we calculate? Using Bayes.

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A')P(A')} = \frac{(.00025)(0.4)}{(.00025)(0.4) + P(.00020)(0.6)} = .4545 \text{ or } 45.45\%$$

Image of a “reverse tree”



Example 2. The answer is .625 or 62.5% -- how do you get there?

5. False Positives and False Negatives

A **false positive** is when a test incorrectly reports that it has found what it is looking for.
A **false negative** is when a test incorrectly reports that it has not found what it is looking for. In other words, the subject has the trait, but the test did not pick it up.

Real world applications are found in blocking “spam”, medical testing, jury verdicts, anti-terrorism methods, credit ratings.

Example: A medical research lab proposes a screening test for a disease. To try out this test, it is given to 100 people, 60 of whom are known to have the disease and 40 of whom are known not to have the disease. A positive test indicates the disease and a negative test indicates no disease. Unfortunately, such medical tests can produce two kinds of errors:

1) A false negative test: For the 60 people who do have the disease, this screening indicates that 2 do not have it.
2) A false positive test: For the 40 people who do not have the disease, this screening test indicates that 10 do have it.

a) Which of the false tests do you think is more serious and why?

b) Incorporate the facts given above into a tree diagram. (Be sure to convert the given integers into probabilities.)

c) Suppose the test is given to a person whose disease status is unknown. If the test result is negative, what is the probability that the person does not have the disease?

a) What is the probability that the diagnostic test gives:

1) A false negative result?

2) A false positive result?

b) What is the probability that:

1) The diagnostic test gives the correct result?

2) A person with a positive diagnostic test has the disease?

3) A person with a negative diagnostic test does not have the disease?

Assignment 3: Please perform a web and/or literature search on the subject of “Bayes” in your area of interest (if you don’t have one, try a search on Bayes with something like “spam” or “guilty” “innocent” or “terrorism”). For example, I found this one pretty quickly:

http://www.techcentralstation.com/022503B.html

After you have gathered a few useful/informative items, please briefly discuss (1) what you found and (2) your understanding of Bayes in these areas. Your understanding does not need to be perfect as this is an ongoing process of learning, just do your best to explain – it does not need to be technical – how Bayes is being used.