Bayes Theorem

John S Butler (TU Dublin)

Bayes Theorem states

$$\Pr(A|B) = \frac{\Pr(B|A)P(A)}{\Pr(B)},$$

where $\Pr(B|A)$ is the probability of B given A and $\Pr(B)$ is the probability of B.

Example: Diagnostic test

The probability that an individual has a rare disease is $\Pr(\text{Disease}) = 0.01$. The probability that a diagnostic test results in a positive (+) test *given you have* the disease is $\Pr(+|\text{Disease}) = 0.95$. On the other hand, the probability that the diagnostic test results in a positive (+) test *given you do not have* the disease is $\Pr(+|\text{No Disease}) = 0.1$. This raises the important question if you are given a positive diagnosis, what is the probability you have the disease $\Pr(\text{Disease}|+)$? From Bayes Theorem we have:

$$\Pr(\text{Disease}|+) = \frac{\Pr(+|\text{Disease})\Pr(\text{Disease})}{\Pr(+)}$$

The probability of a positive test, $\Pr(+)$ is,

$$\Pr(+) = \Pr(+|\text{Disease})\Pr(\text{Disease}) + \Pr(+|\text{No Disease})\Pr(\text{No Disease}),$$

where $\Pr(+|\text{Disease})\Pr(\text{Disease})$ is the Probability of a positive test given the diease among the population of people with the disease and $\Pr(+|\text{No Disease})\Pr(\text{No Disease})$ is the Probability of a positive diagnosis given no diease among the population of people with the diesease,

$$\Pr(+) = 0.1085.$$

Using Bayes Law we can now now calculate the probability of having the disease given a positive diagnosis $\Pr(\text{Disease}|+)$,

$$\Pr(\text{Disease}|+) = \frac{\Pr(+|\text{Disease})\Pr(\text{Disease})}{\Pr(+)} = \frac{0.95 \times 0.01}{0.1085} = 0.0875576.$$

The result means that the rest is only right 8.75% of the time so it is not good test.

This calcualtion can also be done in a simple table format, by assume a population of 10,000, from this we can divide the people into four groups,

Group	+ Diagnosis	- Diagnosis	Total
Disease	95	5	100
No Disease	990	8,910	9,900
Total	1,085	8,915	10,000

From the table we can calculate the same answer,

$$\Pr(\text{Disease}|+) = \frac{\text{Number of people with the disease and + diagnosis}}{\text{Number of people with positive + diagnosis}} = \frac{95}{1085}.$$

Core Concept

This result also gives insight into the core concepts of true positives, true negatives, false positives and false negatives.

Group	+ Diagnosis	- Diagnosis	Total
Disease	true positives	false negative	
No Disease	false positives	true negatives	
Total			

a good test should aim to increase both true positives and true negatives and hence decrease false negatives and false positives.