Assignment - 1

Problem Statement:

A medical test for disease D has outcomes + (positive) and − (negative). We assume that

• the probability for an individual to have the disease is 0.01,

• the probability of a positive test, given that the individual has the disease, is 0.9,

• the probability of a negative test, given that the individual does not have the disease, is 0.9. Compute the probability that an individual has the disease, given that the individual has tested positive. Comment on the quality of the test.

Solution:

To find the probability of a person that the individual has the disease provided the test result of individual is positive. To compute the probability we will use Bayes’ Theorem.

Bayes’ Theorem:

Let A and B be events and assumed P(B) > 0. Then

Where,

P(A) = The initial probability of event A occuring.

P(B) = The initial probability of event B occurring.

P(A|B) = Probability of event A given that event B has occurred.

P(B|A) = Probability of event B occurring given that A has occured.

Information Provided are:

1. The initial probability for an individual to have disease is 0.01 or 1%.

P(Have Disease)=P(D)=0.01

2. The probability of a positive test, given that the individual has the disease, is 0.9 or 90%.

P(Test Positive|Have Disease) = P(Pos|D) =0.9

3. The probability of a negative test, given that the individual does not have the disease, is 0.9

P(Test not Positive|Have no Disease) = P(notPos|notD) = 0.9

We need to compute the probability that an individual has the disease, given that the individual has tested positive.

P(Have Disease|Test Positive) = P(D|Pos)

Assigning the values to Bayes’ Theorem:

Here: P(Pos|D) = 0.9, P(D) = 0.01 and P(Pos) = ?

Calculation:

“P(Pos)” is the total possibility states. In this scenario an individual can be positive in two ways which are true positive result and false positive result.

P(Pos) = P(Pos|D).P(D) + P(Pos|notD).P(notD)

P(Pos|D).P(D): This is the probability of positive result given that individual has the disease. This probability represents positive test from true positive result.

P(Pos|notD).P(notD): This is the probability of positive result given that individual has no disease. This probability represents positive test from false positive result.

P(Pos|notD) = 1 – P(notPos|notD) = 1 – 0.9 = 0.1

P(notD) = 1 – P(D) = 1 – 0.1 = 0.99

P(Pos) = P(Pos|D).P(D) + P(Pos|notD).P(notD)

P(Pos) = (0.9 x 0.01) + (0.1 x 0.99) = 0.108

Assigning all values to following equation:

0.083 or 8.3%

Comment:

1. With high sensitivity of 90% this test correctly identifies the individual who possess the disease.

2. With high specificity of 90% this test also correctly identifies an individual who do not have the disease.

3. The baseline rate of the disease in the population is 1% which is very low. (Per 100 people only 1)

4. Due to very low percentage (1%) of individual having disease, when applied to a large group where very less number will have the disease, there is a chance of projecting false positive.

For Example: If we consider 1000 people, we might get only 10 people (1% of 1000) with disease. Now, only for those 10 people we will find 9 positive results (90%). For rest 990 people who all are not affected by the disease, total 99 individual’s test result will be wrongly positive (10% of 990). For that, 9(true) and 99(false), total positive tested number will be 108. Finally out of 108 positive tests only 9 is found correct which is 8.3%

5. Even though an individual is found positive, a confirmation test is necessary to reduce the false positive result.