Probability practice

Part A. Visitors to your website are asked to answer a single survey question before they get access to the content on the page. Among all of the users, there are two categories: Random Clicker (RC), and Truthful Clicker (TC). There are two possible answers to the survey: yes and no. Random clickers would click either one with equal probability. You are also giving the information that the expected fraction of random clickers is 0.3. After a trial period, you get the following survey results: 65% said Yes and 35% said No. What fraction of people who are truthful clickers answered yes? Hint: use the rule of total probability.

Part B. Imagine a medical test for a disease with the following two attributes:

- The sensitivity is about 0.993. That is, if someone has the disease, there is a probability of 0.993 that they will test positive.
- The specificity is about 0.9999. This means that if someone doesn't have the disease, there is probability of 0.9999 that they will test negative.
- In the general population, incidence of the disease is reasonably rare: about 0.0025% of all people have it (or 0.000025 as a decimal probability).

Suppose someone tests positive. What is the probability that they have the disease?

Solution –

<u>Part A:</u>

<u>Given:</u>

• The total fraction of Random Clickers (RC) is 0.3.

• The total fraction of Truthful Clickers (TC) will then be $1-0.3=0.7$, since there are only these two categories.

• Random Clickers will click "Yes" 50% of the time and "No" 50% of the time since they are clicking randomly.

• The overall fraction of users who clicked "Yes" is 65% (0.65), and the fraction who clicked "No" is 35% (0.35).

<u>Solution:</u>

Using the law of total probability, we can write the fraction of "Yes" answers as a weighted sum of the fractions of "Yes" answers from the two groups (RC and TC):

$0.65 = \text{fraction of RC who said Yes} \times \text{probability of RC} + \text{fraction of TC who said Yes} \times \text{probability of TC}$

Since the fraction of RC who said Yes is 0.5, and the probabilities of RC and TC are 0.3 and 0.7 respectively, we can substitute these values into the equation:

$0.65 = 0.5 \times 0.3 + \text{fraction of TC who said Yes} \times 0.7$

Simplifying, we have:

$0.65 = 0.15 + \text{fraction of TC who said Yes} \times 0.7$

Subtracting 0.15 from both sides, we get:

$0.5 = \text{fraction of TC who said Yes} \times 0.7$

Finally, dividing by 0.7, we find:

$\text{fraction of TC who said Yes} = 0.5/0.7 \approx \mathbf{0.7143}$

So, the fraction of Truthful Clickers who answered "Yes" is approximately **71.43%.**

Part B:

We are tasked with finding out the probability that someone actually has the disease given that they tested positive. This probability can be determined using Bayes' theorem.

Let's define:

• D: The event that someone has the disease.

• +: The event that someone tests positive.

We aim to calculate: $P(D|+)$ Which represents the probability of having the disease given a positive test result.

According to Bayes' theorem: $P(D|+) = P(+|D) \times P(D) / P(+)$

Where:

• $P(+|D)$ is the probability of testing positive given that one has the disease, also known as the sensitivity. This is given as 0.993.

• $P(D)$ is the probability that someone has the disease, which is provided as 0.000025.

We also need to determine P (+), which is the total probability that someone tests positive. This is computed using the total law of probability: $P(+) = P(+|D) \times P(D) + P(+|notD) \times P(notD)$

Where:

• $P(+|notD)$ is the probability of testing positive given that one doesn't have the disease. This is the complement of the specificity, or $1-0.9999=0.00011-0.9999=0.0001$.

• $P(notD)$ is the probability that someone does not have the disease. This is $1-0.000025=0.9999751-0.000025=0.999975$.

Plugging in our known values:

$P(+) = 0.993 \times 0.000025 + 0.0001 \times 0.999975$

$P(+) = 0.000024825 + 0.0000999975 = 0.0001248225$

Now, using our Bayes' theorem formula:

$P(D|+) = 0.993 \times 0.000025 / 0.0001248225$

$P(D|+) \approx$ **0.19888241**

In conclusion, given that someone tests positive for the disease, there's an approximate probability of **19.88%** that they genuinely have the illness.