

Assignment 1

Greeshma Gopinathan (GG33287)

2024-08-18

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.

1. a)

Given :-

Random clickers (RC) fraction = 0.3
Truthful clickers (TC) fraction = 0.7

Random clickers answering Yes = 0.5
 No = 0.5

	Yes	No
RC (0.3)	0.5	0.5
TC (0.7)	65%	35%

Probability of T.C answering "Yes"

Using rule of total probability

$$P(Y) = P(Y|RC) \cdot P(RC) + P(Y|TC) \cdot P(TC)$$
$$0.65 = 0.5 \times 0.3 + P(Y|TC) \times 0.7$$
$$0.65 = 0.15 + P(Y|TC) \cdot 0.7$$
$$\frac{0.65 - 0.15}{0.7} = P(Y|TC)$$
$$P(Y|TC) = 0.714$$

Probability of Truthful answering Yes is
0.714 = 71.4%

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?

1. b)

Given

Sensitivity (TPR) - 0.993
Specificity (TNR) - 0.9999
Prior probability - 0.0025%

Using Bayes Theorem

$$P(\text{Disease} | \text{Positive}) = \frac{P(\text{Positive} | \text{Disease}) \cdot P(\text{Disease})}{P(\text{Positive})}$$

$$P(\text{Positive}) = P(\text{Positive} | \text{Disease}) \cdot P(\text{Disease}) + P(\text{Positive} | \text{No Disease}) \cdot P(\text{No Disease})$$

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

$$= 0.000024825 + 0.0000999975$$

$$= 0.0001248225$$

$$P(\text{Disease} | \text{Positive}) = \frac{0.993 \times 0.000025}{0.0001248225}$$

$$= 0.198\%$$

Probability of someone has disease given positive result is 19.88%