Assignment 1

Tim But, Mehmet Bedirhan Gursoy, Vincentas Ryliskis, group 035

6 November 2023

Task 1

a)

Suppose the following:
- P(D) is the probability of a positive cancer diagnosis
- P(C) is the probability of an individual having cancer

Then, given the information given in exercise 1.3:

$$P(D|C) = 0.95$$

$$P(D|C') = 0.05$$

$$P(C) = 0.004$$

To calculate the probability of a random person being given a positive cancer diagnosis, the following must be calculated:

$$P(D) = P(D|C) * P(C) + P(D|C') * P(C')$$
$$= 0.95 * 0.04 + 0.05 * 0.996$$
$$= 0.0536$$

This differs from probability we are asked to calculate in the Exercise 1.3, as there we are asked to calculate the probability of someone having cancer given a cancer diagnosis, or: P(C|D). These probabilities refer to different events, as the latter already assumes P(D) as being true.

b)

As mentioned above, we need to calculate P(C|D). We can calculate this value using Bayes' Theorem

$$P(C|D) = \frac{P(D|C) * P(C)}{P(D)} = \frac{0.95 * 0.004}{0.0536} \approx 0.071$$

c)

The two events that a person has cancer and that the test is positive are dependent. This is shown by the fact that probabilities P(C|D) and P(C) are different, as shown below (based on calculations done in parts a) and b)):
- $P(C|D) = 0.071$
- $p(C) = 0.004$
The fact that P(C|D) is larger than P(C) shows that a cancer diagnosis leads to an increased risk in cancer, when compared to the population average.

Task 2

a)

b)

c)

d)

e)

Task 3

a)

(i)

(ii)

b)

(i)

(ii)

(iii)

(iv)

c)

(i)

(ii)

Task 4

a)

b)

c)

d)