homework_01

August 31, 2020

1 Homework 1 Solution

1.1 Problem 1

This exercise demonstrates that probability theory is actually an extension of logic. Assume that you know that "A implies B." That is, your prior information is:

$$I = \{A \implies B\}.$$

Please answer the following questions in the space provided:

A. (4 points) $p(AB|I) = p(A|I)$. **Proof:**

Using the product rule, we get:

$$p(AB|I) = p(B|AI)p(A|I).$$

Now, notice that $p(B|AI) = 1$ because when A is true B must be true according to information I. From this, the result follows.

B. If $p(A|I) = 1$, then $p(B|I) = 1$. **Proof:**

If $p(A|I) = 1$ then $p(\neg A|I) = 0$.

By the sum rule:

$$p(B|I) = \underbrace{p(B|AI)}_{=1}p(A|I) + p(B|\neg A|I)\underbrace{p(\neg A|I)}_{=0}.$$

Thus, if $p(A|I) = 1, p(B|I) = 1$. What does this mean? Is it consistent with the fact that "A implies B?" Think about it.

C. If $p(B|I) = 0$, then $p(A|I) = 0$. **Proof:**

We have from the product rule:

$$p(AB|I) = p(A|BI)p(B|I),$$

and

$$p(AB|I) = p(B|AI)p(A|I).$$

Therefore, we get:

$$p(A|BI)p(B|I) = p(B|AI)p(A|I) = p(A|I) = 0.$$

So, if $p(A|BI) > 0$, then we get that $p(B|I) = 0$. This is the probabilistic way of saying that "not B implies not A."

But why is $p(A|BI) > 0$? Assume that this is not true. Take $p(A|BI) = 0$ and that $p(B|I) > 0$. What does this mean? It means that A is definitely false if B is true and that there is some probability that B is true. But because "A implies B," if there is some probability that B is true, there must be some probability that A is true given B. This is a contradiction.

D. B and C show that probability theory is consistent with Aristotelian logic. Now, you will discover how it extends it. Show that if B is true, then A becomes more plausible, i.e.

$$p(A|BI) \geq p(A|I).$$

Proof:

From the product rule, we have:

$$p(A|BI)\underbrace{p(B|I)}_{\leq 1} = \underbrace{p(B|AI)}_{=1}p(A|I) = p(A|I).$$

From this we get:

$$p(A|BI) \geq p(A|I).$$

E. Give at least two examples of D that apply to various scientific fields. To get you started, here are two examples:

- A: It is raining. B: There are clouds in the sky. Clearly, $A \implies B$. D tells us that if there are clouds in the sky, raining becomes more plausible.
- A: General relativity. B: Light is deflected in the presence of massive bodies. Here $A \implies B$. Observing that B is true makes A more plausible.

Answer:

There is an infinity of correct answers here.

- A: $f(x)$ is differentiable at x_0. B: $f(x)$ is continuous at x_0.

...

F. Show that if A is false, then B becomes less plausible, i.e.:

$$p(B|\neg AI) \leq p(B|I).$$

Proof:

$$\begin{aligned}
p(B|\neg AI) &= \frac{p(\neg A|BI)p(B|I)}{p(\neg A|I)}, \\
&= \frac{(1-p(A|BI))p(B|I)}{p(\neg A|I)}, \\
&\leq \frac{(1-p(A|I))p(B|I)}{p(\neg A|I)}, \text{(from the result in part (D))}, \\
&= \frac{p(\neg A|I)p(B|I)}{p(\neg A|I)}, \\
&= p(B|I).
\end{aligned}$$

G. Can you think of an example of scientific reasoning that involves F? For example: A: It is raining. B: There are clouds in the sky. F tells us that if it is not raining, then it is less plausible that there are clouds in the sky. **Answer:**

Again an infinity of answers here. Any example from part (E) would work.

H. Do D and F contradict Karl Popper's principle of falsification, "A theory in the empirical sciences can never be proven, but it can be falsified, meaning that it can and should be scrutinized by decisive experiments." **Answer:**

No it doesn't.

D and F are consistent with Popper's principle of falsification. Suppose A is a scientific theory and B is the prediction from the theory, i.e, $A \implies B$.

(i) If prediction B matches empirical reality, it increases our belief that theory A is true. This is the statement of part D. This statement does not violate Popper's principle. It doesn't prove the theory (which is forbidden by Popper's principle). It simply makes its truth more likely.

(ii) If the theory A is incorrect, the plaussibility of the prediction B decreases. This is the statement of part F.

1.2 Problem 2

Disclaimer: This example is a modified version of the one found in a 2013 lecture on Bayesian Scientific Computing taught by Prof. Nicholas Zabaras. I am not sure where the original problem is coming from.

We are tasked with assessing the usefulness of a tuberculosis test. The prior information I is:

> The percentage of the population infected by tuberculosis is 0.4%. We have run several experiments and determined that: + If a tested patient has the disease, then 80% of the time the test comes out positive. + If a tested patient does not have the disease, then 90% of the time the test comes out negative.

To facilitate your analysis, consider the following logical sentences concerning a patient:

A: The patient is tested and the test is positive.

B: The patient has tuberculosis.

A. Find the probability that the patient has tuberculosis (before looking at the result of the test), i.e., $p(B|I)$. This is known as the base rate or the prior probability. **Answer:**

From the question details we get the base rate: $p(B) = 0.004$.

NOTE: The conditioning on I is implicit.

B. Find the probability that the test is positive given that the patient has tuberculosis, i.e., $p(A|B, I)$. **Answer:**

From the question details we get: $p(A|B) = 0.8$.

C. Find the probability that the test is positive given that the patient does not have tuberculosis, i.e., $p(A|\neg B, I)$. **Answer:**

From the question details we get the probability of the patient testing negative if he/she does not have Tuberculosis: $p(\neg A|\neg B) = 0.9$.

By the sum rule, the probability of the patient testing positive given that the patient does not have Tuberculosis is: $p(A|\neg B) = 1 - 0.9 = 0.1$.

D. Find the probability that a patient that tested positive has tuberculosis, i.e., $p(B|A, I)$. **Answer:**

From Bayes' rule: $p(B|A) = \frac{p(B)p(A|B)}{p(A)}$.

The normalization constant $p(A) = p(A|B)p(B) + p(A|\neg B)p(\neg B) = 0.1028$.

Thus, the probability of a patient having tuberculosis given that the patient tested positive is:

$$p(B|A) = \frac{0.004 \times 0.8}{0.1028} = 0.031128.$$

E. Find the probability that a patient that tested negative has tuberculosis, i.e., $p(B|\neg A, I)$. Does the test change our prior state of knowledge about about the patient? Is the test useful? **Answer:**

From Bayes' rule: $p(B|\neg A) = \frac{p(B)p(\neg A|B)}{p(\neg A)} = \frac{0.004 \times 0.2}{1-0.1028} = 0.000891 \approx 0.09\%$.

F. What would a good test look like? Find values for

$$p(A|B, I) = p(\text{test is positive}|\text{has tuberculosis}, I),$$

and

$$p(A|\neg B, I) = p(\text{test is positive}|\text{does not have tuberculosis}, I),$$

so that

$$p(B|A, I) = p(\text{has tuberculosis}|\text{test is positive}, I) = 0.99.$$

There are more than one solutions. How would you pick a good one? Thinking in this way can help you set goals if you work in R&D. If you have time, try to figure out whether or not there exists such an accurate test for tuberculosis **Answer:**

Denote $p(A|B)$ and $p(A|\neg B)$ with x and y respectively.

Bayes' rule: $p(B|A) = \frac{p(B)p(A|B)}{p(B)p(A|B)+p(\neg B)p(A|\neg B)} = \frac{0.004x}{0.004x+0.996y}$.

Rearranging the terms, we get: $0.00004x = 0.986y$.

There are several criteria one could choose for a good test. One such criteria could be that the patient tests positive with very high probability if he/she has the disease. Ideally, $p(A|B) = x = 1$. Then $y = p(A|\neg B) = 4.058 \times 10^{-5}$.