Homework 1

Due 9/1/2020.

References

- Lectures 1-2 (inclusive).

Instructions

- Type your name and email in the "Student details" section below.
- Develop the code and generate the figures you need to solve the problems using this notebook.
- For the answers that require a mathematical proof or derivation you can either:
  - Type the answer using the built-in latex capabilities. In this case, simply export the notebook as a pdf and upload it on gradescope; or
  - You can print the notebook (after you are done with all the code), write your answers by hand, scan, turn your response to a single pdf, and upload on gradescope.
- The total homework points are 100. Please note that the problems are not weighed equally.

Note: Please match all the pages corresponding to each of the questions when you submit on gradescope.

In [2]:
```
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_context('paper')
import numpy as np
```

Student details

- **First Name:** Kiara
- **Last Name:** Koeppen
- **Email:** kkoeppen@purdue.edu

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:

From the product rule we know that:

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

Since we know that A implies B we can say "If A is true then B MUST be True". In otherwords, the probablility of B being true given that A is true is 1 because whenever A is true we know B has to be true (we're 100% sure that if A is true B will also be true). This means that:

$$p(B|A,I) = 1$$

because the probability of B being true given A is true is 1.

now we have the expression:

$$p(AB|I) = (1)p(A|I)$$

Proving that:

$$p(AB|I) = P(A|I)$$

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

Now we have 2 "conditions":

1. A implies B
2. p(A|I) = 1

The first condition tells us

$$p(B|A, I) = 1$$

The second condition tells us that A is always true. Knowing that A is always true, we know that B must also always be true because if A is true B must be true given condition 1. Since B must be true given that we know A is true, the probability of A being true given that B is true is 1:

$$p(A|B, I) = 1$$

From the product rule we have these two equations:

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

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

Combining them gives us this:

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

From condition 2 we know:

$$p(A|B, I) = 1$$

$$p(B|A, I) = 1$$

We are left with:

$$p(B|I) = p(A|I)$$

Since we know p(A|I)=1 we can say:

$$p(B|I) = 1$$

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

We now have 2 new "conditions".

1. A implies B
2. p(B|I)=0

In other words, condition 1 tells us that P(B|A,I)=1 because the probability of B being true given that A is true is 1 since if A is true B must be true. Condition 2 tells us that the probability of B being true is 0. Since we know B is always "false" we know that A can never be true (given condition 1).

From the product rule we have these two equations:

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

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

Combining them gives us this:

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

We now know:

$$p(B|A,I) = 1$$

$$p(B|I) = 0$$

$$(0)p(A|B,I) = (1)p(A|I)$$

$$0 = p(A|I)$$

$$p(A|I) = 0$$

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:

We have 2 conditions:

1. A implies B
2. B is true (a non-zero number)

We know:

p(B|A,I)=1 (given condition 1)

0<p(B|I)≤1 (given condition 2)

From the product rule we have these two equations:

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

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

Combining them gives us this:

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

$p(A|B, I)(0 < p(B|I) \leq 1) = (1)p(A|I)$

Since we are equating these two, in order for the expression to equate, p(A|B,I) must be greater than or equal to p(A|I) because we are multiplying p(A|I) by 1 and p(A|B,I) by a number less than or equal to 1. In order for the left side of the equation to equal the right side of the equation, p(A|B,I) must be a larger number (when p(B|I) is a decimal) and it will equal p(A|I) when p(B|I) is 1.

In other words:

p(A|B,I)(decimal number) = (1)p(A|I) then p(A|B,I) must be greater than p(A|I) to equate the left and right sides of te equation.

However, when p(B|I)=1, then p(A|B,I)=p(A|I).

So we are left with

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

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:

- A: Global warming due to ozone depletion

  B: Increase of CO2 in the atmosphere

  If we are observing that there is an increase of CO2 in the atmosphere then global warming due to ozone depletion is more plausible.

- A: People are ignoring CDC guidlines for Covid

  B: We are seeing more positive cases of Covid

  If we are seeing more positive covid cases, then it is more plausible that people are ignoring the CDC's guidlines for Covid

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

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

Proof: We have 2 "conditions"

1. A implies B (meaning p(B|A,I)=1)
2. A is false (since we know A is false this means we don't know anything about B, B can be true or false)

From the product rule we have these two equations: $p(AB|I) = p(A|B,I)p(B|I)$

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

Rearranging this we have

$p(B|\neg A,I) = \frac{p(\neg AB|I)}{p(\neg A|I)} = \frac{p(\neg A|B,I)p(B|I)}{p(\neg A|I)}$

Expanding this by using the obvious rule we get:

$\frac{(1-p(A|B,I))(p(B|I))}{1-p(A|I)} = \frac{1-p(A|B,I)p(B|I)}{(1-p(A|I))p(B|I)}$

After simplifying this we get:

$p(B|\neg A,I) = \frac{1-p(A|B,I)}{1-p(A|I)}p(B|I)$

From 1E we proved that: $p(A|BI) \geq p(A|I)$ so using that we can say

$-p(A|BI) \leq -p(A|I)$ (multiply everything by -1 on both sides so sign flip)

$1 - p(A|BI) \leq 1 - p(A|I)$ (add 1 to both sides, still the same equation)

$\frac{1-p(A|BI)}{1-p(A|I)} \leq 1$ (Bring all the variables on one side)

$\frac{1-p(A|BI)}{1-p(A|I)}p(B|I) \leq p(B|I)$ (multiply both sides by p(B|I))

Now we know that:

$p(B|\neg A,I) = \frac{1-p(A|B,I)}{1-p(A|I)}p(B|I)$

$\frac{1-p(A|BI)}{1-p(A|I)}p(B|I) \leq p(B|I)$

We can no say that:

$p(B|\neg A,I) \leq p(B|I)$

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:

A: Manufacturing tolerance specs are off

B: Defective part

If manufacturing specs are correct (not off) then it is less plausible that we're going to have a defective part

H. Do D and F contradict Karl Popper's [principle of falsification (https://en.wikipedia.org/wiki/Falsifiability)](https://en.wikipedia.org/wiki/Falsifiability), "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:

D does contradict Karl Popper's statement because it is saying "If B is true then A becomes more true". Popper's is saying that we can never prove something is "more true" we can only disprove it. F do not contradict this statement because it's saying if A is false B becomes "more false". This is what Karl Popper is stating since we are saying something is "more false" (we are denying B to be true by proving it is "more false").

Essentially D is contradictory because Popper is saying nothing can be proven to be "more true" and F is not contradictory because it is showing that something is "more false".

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:

If 0.4% of the population has tuberculosis, then we know that the probability of a patient having tuberculosis is 0.004.

Therefore

$$p(B|I) = 0.004$$

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

If we get a positive test result, 80% of the time it will be true (they actually have tuberculosis) and 10% of the times it will be false (they don't have tuberculosis).

p(A|B,I) is the probability of a test being positive, given that patient has tuberculous.

Meaning p(A|B,I)=0.8 because 80% of the time, the test comes out positive given the patient has tuberculous.
p(B|A,I) represents the probability that the patient has tuberculous given the test is positive.

$p(A|B, I) = 0.8$

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:

If the patient does not have tuberculosis then 90% of the time it will come back with a negative test result, however that means that 10% of the time the test will come back as a false positive.

Show in equations:

$p(\neg A|\neg B, I)$= 0.9 (the probability that the patient does not test positive given that the patient does not have tuberculosis)

Obvious Rule:

$p(\neg A|\neg B, I) + p(A|\neg B, I) = 1$

$p(A|\neg B, I)$= 0.1

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

We know:

$p(B|I) = 0.004$

$p(A|B,I) = 0.8$

$p(A|\neg B,I) = 0.1$

$p(\neg B|I) = 0.996$ (given the obvious rule)

Sum Rule Given that A and B are mutually exclusive events:

$p(A|I) = p(A|B,I)p(B|I) + p(A|\neg B,I)p(\neg B|I)$

$p(A|I) = (0.8)(0.004) + (0.1)(0.996) = 0.1028$

From the product rule we have these two equations:

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

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

Combining these gives us:

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

Since we now have the equation for p(A|I) we can plug that in and solve

$(0.8)(0.004) = p(B|A,I)(0.1028)$

$p(B|A,I) = 0.031$

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:

We know:

$p(B|I) = 0.004$

$p(A|B, I) = 0.8$

$p(A|\neg B, I) = 0.1$

$p(\neg B|I) = 0.996$ (given the obvious rule)

$p(B|A, I) = 0.031$ (previous problem)

$p(A|I) = 0.1028$

$p(\neg A|I) = 0.8972$ (given obvious rule)

$p(\neg B|A, I) = 0.969$ (given obvious rule)

Sum Rule Given that A and $\neg A$ are mutually exclusive events:

$p(B|I) = p(B|A, I)p(A|I) + p(B|\neg A, I)p(\neg A|I)$

$(0.004) = (0.031)(0.1028) + p(B|\neg A, I)(0.8972)$

$p(B|\neg A, I) = 0.000906$

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:

Knowns:

$p(B|I) = 0.004$

$p(\neg B|I) = 0.996$ (given the obvious rule)

$p(B|A,I) = 0.99$

If we want a good test, p(A|B,I) will be as close to 1 as possible. This is because ideally any time someone has tuberculosis they will test positive for it. On the other side, $p(A|\neg B, I)$ will be as close to 0 as possible. This is because if the patient does not have tuberculosis then they will NOT test positive.

To relate these two variables we wil define a constant, k

$p(A|B,I) = kp(A|\neg B,I)$

Sum Rule Given that A and B are mutually exclusive events:

$p(A|I) = p(A|B,I)p(B|I) + p(A|\neg B,I)p(\neg B|I)$

From the product rule we have these two equations: $p(AB|I) = p(A|B,I)p(B|I)$

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

Eauating these gives us:

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

Plugging in the equation for p(A|I) gives us:

$p(A|B,I)p(B|I) = p(B|A,I)((p(A|B,I)p(B|I) + p(A|\neg B,I)p(\neg B|I))$

$k * p(A|\neg B,I)(0.004) = (0.99)((k * p(A|\neg B,I)(0.004) + p(A|\neg B,I)(0.996))$

$(\frac{0.004}{0.99} - 0.004)k * p(A|\neg B,I) = 0.996p(A|\neg B,I)$

$k = 24651$

If we had a "bad" test and said that p(A|B,I)=0.5 then we'd show that $p(A|\neg B, I)$=0.0000203

If we use our example p(A|B,I)=0.8, then $p(A|\neg B, I) = 0.000040161$.

If we have a perfect test, p(A|B,I)=1, then we get $p(A|\neg B, I)$=0.000032453

This shows that as the accuracy of the test increases, the probability that you will have a "false positive" increases as well given p(B|A,I)=0.99. I started the examples with p(A|B,I)=0.5 and ended it with a perfect test, p(A|B,I)=1. The probability of getting a positive tuberculosis test while not having tuberculosis increased as well. However, since the probability of getting a "false positive" is so small to begin with, the increase is minimal.

Since the "false positives" change so minimally with p(B|A,I) being 0.99 then it really depends on how accurate we want our p(A|B,I) to be. I believe a "good" test would be at least 90% accurate.