MTH794P Probability and Statistics for Data Analytics — Problem Sheet 3

Kavit Tolia

October 16, 2025

Problem 1.

I have three coins in my pocket. Two of them are ordinary fair coins, the third has Heads on both sides.

(a) I take a random coin from my pocket and toss it. What is the probability that it comes up Heads?

(b) Given that it did come up Heads, what is the conditional probability that it is the double-headed coin?

(c) Given that it did come up Heads, what is the conditional probability that a second toss of the same coin also comes up Heads?

(d) Given that two tosses of the same coin both come up Heads, what is the conditional probability that it is the double-headed coin? [Before doing this think do you expect the answer to be larger or smaller than the answer to part (b)? Why?]

(e) Suppose that 100 tosses of the same coin show a Head every time. Without doing any more calculations, say roughly what you expect the conditional probability that it is the double-headed coin to be? How would you explain this in non-mathematical terms?

Solution.

(a) Let $\mathbb{P}(H)$ be the probability of the coin showing Heads, $\mathbb{P}(F)$ be the probability that it is a fair coin and $\mathbb{P}(B)$ be the probability that it is a biased coin. Then,

$$\mathbb{P}(H) = \mathbb{P}(H \mid F)\,\mathbb{P}(F) + \mathbb{P}(H \mid B)\,\mathbb{P}(B) = \left(\frac{1}{2}\right)\left(\frac{2}{3}\right) + (1)\left(\frac{1}{3}\right) = \frac{2}{3}$$

(b) Now we want to work out $\mathbb{P}(B|H)$:

$$\mathbb{P}(B \mid H) = \frac{\mathbb{P}(H \mid B)\mathbb{P}(B)}{\mathbb{P}(H)} = \frac{(1)\left(\frac{1}{3}\right)}{\frac{2}{3}} = \frac{1}{2}$$

(c) Now let's say H_i is the i^{th} toss, then we want to calculate $\mathbb{P}(H_2 = H \mid H_1 = H)$:

$$\mathbb{P}(H_2 = H \mid H_1 = H) = \frac{\mathbb{P}(H_2 = H \cap H_1 = H)}{\mathbb{P}(H_1 = H)} = \frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4}$$

(d) I expect this probability to be higher than (b), as I'm more certain of the fact that it is now a baised coin.

$$\mathbb{P}(B \mid HH) = \frac{\mathbb{P}(HH \mid B)\mathbb{P}(B)}{\mathbb{P}(HH)} = \frac{(1)\left(\frac{1}{3}\right)}{\frac{1}{2}} = \frac{2}{3}$$

(e) I would expect this conditional probability to be close to 1, as it is highly unlikely that the coin is a fair coin. If it was a fair coin, the chances of getting 100 Heads in a row is close to impossible. The only plausible explanation is that it is the double-headed coin.

Problem 2.

A certain medical condition affect 1% of the population. A new AI tool for detecting this condition from a scan has been developed. It has a 95% success rate at correctly detecting that a person with the condition has it, and only a 2% chance of incorrectly deciding that a healthy person has the condition. A randomly chosen person from the population undertakes this test and the test shows positive.

(a) What is the probability that they do have the condition?

(b) A politician suggests that this test could be used for a national screening programme. What insight into the possible disadvantages of doing this does the calculation of part (a) provide? How could these disadvantages be mitigated?

Solution.

(a) Let's define some outcomes. Let H denote a person being healthy and P denote a positive test result. Then,

$$\mathbb{P}(H^c \mid P) = \frac{\mathbb{P}(P \mid H^c)\mathbb{P}(H^c)}{\mathbb{P}(P)} = \frac{(95\%)\,(1\%)}{(95\%)\,(1\%) + (2\%)\,(99\%)} = 32.5\%$$

(b) There is a major disadvantage here in that you have less than 1 in 3 chance of a person actually having the condition if they tested positive. This seems odd given the test's high accuracy. But this happens due to the low rate of the condition in the general population.

We can mitigate this by either:

- We can use the test as a first screen, followed by more robust tests for those who are positive.
- Only screen high-risk groups, where we know the prevalence to be higher, significantly increasing the probability that a positive test is indeed someone with the condition.

Problem 3.

Suppose that A and B are events with $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$ and that $\mathbb{P}(A \mid B) > \mathbb{P}(A)$.

(a) What can you say about $\mathbb{P}(B \mid A)$?

(b) How do you explain the relationship between events A and B which satisfy this property in non-mathematical language?

Solution.

(a) Let us write this out:

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\mathbb{P}(B)}{\mathbb{P}(A)} > \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(A)} = \mathbb{P}(B)$$

So, we have $\mathbb{P}(B \mid A) > \mathbb{P}(B)$

(b) This tells us that the occurance of one event makes the other more likely to happen. In laymen's terms, A and B are positively linked, as they occur together more often.

Problem 4.

Let T be the random variable giving the time (in minutes) between consecutive customers arriving in a shop. Suppose that $T \sim \text{Exp}(0.5)$. Each customer spends 5 minutes in the shop and then leaves. The first customer of the day has just entered the shop.

(a) How do your answers to (c) and (d) change if the parameter of T changes?

Solution.

(a) We are looking for the probability that $T > 5$:

$$\mathbb{P}(T > 5) = 1 - \mathbb{P}(T \leq 5) = 1 - F_T(5) = 1 - (1 - e^{-\lambda(5)}) = e^{(-2.5)} = 0.08$$

(b) $\mathbb{E}(T) = 1/\lambda = 1/0.5 = 2$

Problem 5.

Read this article from mathematical epidemiologist Adam Kucharski's blog:
`https://kucharski.substack.com/p/small-hallucinations-big-problems`

(a) Using what you learn in this module, redo this analysis in the case that the probability of the unusual event we are looking for is p (rather than the 1 in 1000 that the article uses).

(b) Do you think the author did a good job at explaining the mathematics to a non-mathematical audience?

(c) Look at the Student Forum on the module QMPlus page and make a comment (which could be about your answer to parts (a) or (b) of this question or something else) on the discussion thread about this article.

Solution.

(a) If the probability of the unusual even is p, then let us define E as the event occurring and F as it being flagged by the LLM. Then,

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(F \mid E)\,\mathbb{P}(E)}{\mathbb{P}(F \mid E)\,\mathbb{P}(E) + \mathbb{P}(F \mid E^c)\,\mathbb{P}(E^c)} = \frac{0.99p}{0.99p + 0.01(1-p)}$$

We can use this formula to understand how the rarity of an event impacts the probability of a flag being correct:

p	0.1%	1.0%	10.0%
$\mathbb{P}(E \mid F)$	0.01	0.01	1.00

Table 1: How p impacts the flag being correct.