# Exercises on Discrete Mathematics and Probability

Samson Abramsky

## Question 1

Suppose that $X$ is a set with $m$ elements, and $Y$ is a set with $n$ elements.

(a) How may elements does $Y^X$ have?

(b) How many injective functions from $X$ to $Y$ are there?

(c) Two football teams are playing each other in the World Cup. Show that at least four players from the two teams must have been born on the same day of the week.

## Question 2

A **partition** of a set $X$ is a pairwise disjoint family of non-empty sets $\{X_i\}_{i \in I}$ such that $\bigcup_{i \in I} X_i = X$. A $k$-**partition** of $X$ is a partition $\{X_1, \ldots, X_k\}$.
Show that there is a one-one correspondence between:

1. surjective functions $s : X \to \{1, \ldots, k\}$
2. $k$-partitions of $X$.

## Question 3

Let $\mathbf{1} := \{0\}$.

(a) How many functions $X \to \mathbf{1}$ are there?

(b) How many relations $R : X \multimap \mathbf{1}$ are there?

## Question 4

If we have functions $s : X \to Y$ and $r : Y \to X$ such that $r \circ s = \mathsf{id}_X$, then $s$ is called a **section**, and $r$ a **retraction**.

(a) Show that a function is injective if and only if it is a section (for some retraction).

(b) Show that a function is surjective if and only if it is a retraction (for some section).

**Note** Showing that a surjective function is a retraction, for general (infinite) sets, is equivalent to the **Axiom of Choice** in set theory.

## Question 5

An HIV test gives a positive result with probability 98% when the patient is indeed affected by HIV, while it gives a negative result with 99% probability when the patient is not affected by HIV. If a patient is drawn at random from a population in which 0.1% of individuals are affected by HIV and he is found positive, what is the probability that he is indeed affected by HIV?

## Question 6

There are two urns containing colored balls. The first urn contains 50 red balls and 50 blue balls. The second urn contains 30 red balls and 70 blue balls. One of the two urns is randomly chosen (both urns have probability 50% of being chosen) and then a ball is drawn at random from one of the two urns. If a red ball is drawn, what is the probability that it comes from the first urn?

## Question 7

Alice has two coins in her pocket, a fair coin (head on one side and tail on the other side) and a two-headed coin. She picks one at random from her pocket, tosses it and obtains head. What is the probability that she flipped the fair coin?

## Question 8

Prove the following properties of conditional probabilities:

(a) If $A_1 \cap B$ and $A_2 \cap B$ are disjoint, then:

$$p(A_1 \cup A_2|B) \;=\; p(A_1|B) + p(A_2|B).$$

(b) $p(A^c|B) \;=\; 1 - p(A|B)$.

(c) If $p(A) \geq p(B)$, then $p(A|B) \geq p(B|A)$.

## Question 9

Let $S$ be the set of all applicants for vacant jobs in a university in a given period. All applicants are female ($F$) or male ($F^c$). Every applicant applies either to Philosophy ($P$) or to Computer Science ($P^c$). Let $H$ be the set of applicants actually hired. Suppose that

$$p(H|F \cap P) > p(H|F^c \cap P)$$

*i.e.* the probability that an arbitrarily chosen female applicant to Philosophy is hired is greater than the probability that an arbitrarily chosen male applicant to Philosophy is hired.

Suppose that the same situation holds with respect to applicants to Computer Science, *i.e.*

$$p(H|F \cap P^c) > p(H|F^c \cap P^c).$$

Must it follow that the overall probability that a female applicant will be hired is greater than that a male applicant will be hired, *i.e.*

$$p(H|F) > p(H|F^c)?$$

It seems that this must be true, but in fact, this is **not the case**. This is known as **Simpson's paradox** (although it is not a paradox!).

To see this, consider the following example. $S$ contains 26 applicants, 13 male and 13 female. Eight women apply to Computer Science, of whom two are hired, and five men apply to Computer Science, of whom one is hired. Five women apply to Philosophy, of whom four are hired, while eight men apply, and six are hired.

Check that this provides a counterexample, where the first two inequalities hold, but the third does not.

**Answer 5**

The given information can be summarised as follows:

$$
\begin{aligned}
p(\text{positive}|\text{HIV}) &= 0.98 \\
p(\text{positive}|\text{NO HIV}) &= 1 - 0.99 = 0.01 \\
p(\text{HIV}) &= 0.001 \\
p(\text{NO HIV}) &= 0.999
\end{aligned}
$$

The probability of being found positive can be derived using the law of total probability:

$$
\begin{aligned}
p(\text{positive}) &= p(\text{positive}|\text{HIV})p(\text{HIV}) + p(\text{positive}|\text{NO HIV})p(\text{NO HIV}) \\
&= 0.98 \cdot 0.001 + 0.01 \cdot 0.999 \\
&= 0.00098 + 0.00999 \\
&= 0.01097
\end{aligned}
$$

So, by Bayes' Theorem:

$$
\begin{aligned}
p(\text{HIV}|\text{positive}) &= \frac{p(\text{positive}|\text{HIV})p(\text{HIV})}{p(\text{positive}} \\
&= \frac{0.98 \cdot 0.001}{0.01097} \\
&= \frac{0.00098}{0.01097} \simeq 0.08933.
\end{aligned}
$$

So despite the conditional accuracy of the test, the unconditional probability of being affected by HIV when found positive is less than 10%.