Probability and Statistics for Data Analytics – 2025/26
Problem Sheet 1

The questions on this sheet are based on the material on events and discrete random variables from Week 1 lectures. This sheet is not for assessment. We will discuss selected questions from it in the Week 2 seminar.

---

1. You toss a fair coin until repeatedly, recording the outcome of each toss. You continue until you have seen either two Heads or three Tails and then you stop.

   (a) Write down the sample space for this experiment.

   (b) Would it be reasonable to assume that all outcomes of the sample space are equally likely? Explain.

   (c) Write down the event "You toss the coin exactly four times" as a subset of the sample space.

   (d) Write down in words, a random variable related to this experiment. What values does it take?

2. Three runners, Amy, Bea and Cate, take part in a race. The order in which they finish is recorded.

   (a) Write down the sample space for this experiment.

   (b) Write down the event Amy finishes ahead of Cate as a set.

   (c) Write down another event both in words and as a set.

   (d) Suppose you had n runners rather than 3. What is the sample space now and how many elements does it contain?

3. Let A, B, C be events with

$$\mathbb{P}(A) = 0.7; \quad \mathbb{P}(B) = 0.6; \quad \mathbb{P}(C) = 0.5; \quad \mathbb{P}(A \cap B) = 0.4$$

$$\mathbb{P}(A \cap C) = 0.3; \quad \mathbb{P}(B \cap C) = 0.3; \quad \mathbb{P}(A \cap B \cap C) = 0.2.$$

Calculate the following:

(a) $\mathbb{P}(A \cup B)$

(b) $\mathbb{P}(A \setminus B)$

(c) The probability that neither of A and B occur.

(d) $\mathbb{P}(A \cup B \cup C)$

(e) The probability that exactly two of A, B and C occur.

4. You have a choice of picking one of two processes each of which produces a random number. You would prefer a higher numbers as the outcome. If you pick the first process the outcome is random variable X; if you pick the second process the outcome is random variable Y. The random variables X and Y have pmfs as follows:

n	0	1	2	3
$\mathbb{P}(X = n)$	0.1	0.2	0.3	0.4

n	0	1	2	3
$\mathbb{P}(Y = n)$	0.2	0.2	0.1	0.5

(a) Which procedure would you pick?

(b) Give some reasons to justify your answer.

(c) Suppose that these two processes represent two possible medical procedures to treat a condition, with the numerical outcome being the quality of life of a patient after the treatment. How does this extra context change the way you think about the choice between these procedures?

5.

(a) Consider the following game. You pick an amount of money n. Then we toss a fair coin. If it comes up Heads, I give you £n; if it comes up Tails, you give me £n. We repeat the game (you can choose a different amount each time) until you decide to stop.

You decide to adopt the following strategy:

- On the first go stake £1.
- If you win stop.
- If you lose then double your stake on the next game.
- Repeat this (doubling your stake after each loss) until you win.

You argue as follows:

- However many turns the game lasts you will win £1. For example, if it takes 3 turns before the coin comes up Heads, your total gain in pounds is $-1-2+4=1$. So you are guaranteed to make £1.
- The game shouldn't last too long. If we let T be the number of times the coin is tossed then you remember (or look up) that T has a Geometric distribution with parameter $1/2$ so $\mathbb{E}(T)=2$.
- Since in round r you stake £2^{r-1}, the expectation of the amount of money you expect to risk in the final round is $\mathbb{E}(2^{T-1})$ which is also small.

What do you think of each stage of this argument? Is this a sensible strategy to use?

(b) Here is another coin game. We toss a coin until the first time it comes up Heads. Suppose this is on toss number N. If N is even, I pay you £2^N; if N is odd, you pay me £2^N. Let W be your total winnings (which could be be negative if you end up paying me!). Find the pmf of the random variable W. What can you say about $\mathbb{E}(W)$?

Robert Johnson
r.johnson@qmul.ac.uk