Part IA: Mathematics for Natural Sciences B
Examples Sheet 8: Probability

Please send all comments and corrections to jmm232@cam.ac.uk.

Questions marked with a ($*$) are difficult and should not be attempted at the expense of the other questions.

Sample spaces and events

1. In an experiment, two fair four-sided dice are rolled. We define:

$$S_1 = \{(i,j) : i \text{ is the result of the first die, } j \text{ is the result of the second die}\},$$
$$S_2 = \{\text{the sum of the results is odd, the sum of the results is even}\},$$
$$S_3 = \{\text{the sum of the results is prime, the first die shows } 1, \text{ the first die shows } 2\}.$$

   Which of S_1, S_2, S_3 are valid sample spaces for the experiment?

2. Given a (discrete) sample space S, define an *event*. Write down in set notation:
   (a) a sample space for the result of a 12-sided die role, the event corresponding to getting a three, the event corresponding to getting an even result, and the event corresponding to getting a prime result;
   (b) a sample space for the result of flipping three coins, the event corresponding to getting all tails, the event corresponding to getting an even number of tails, and the event corresponding to getting more heads than tails.

3. Suppose that S is a sample space and A, B, C are events. By drawing appropriate diagrams, show that:

   (a) $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$, (b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

   where in (a) the bar denotes the complement of a set in S.

4. State the set-theoretic definition that two events are *mutually exclusive*. Show that if A, B, C are events with A, B mutually exclusive, then $A \cap (B \cup C) = A \cap C$.

Probability measures

5. Let S be a (discrete) sample space, and let $\mathcal{F}$ be the set of all associated events. What are the three basic axioms that a probability measure $\mathbb{P} : \mathcal{F} \to [0,1]$ must satisfy?[1]

   Now suppose that $S = \{\omega_1, \omega_2, \omega_3\}$ is a sample space containing three outcomes.
   (a) Write down the set of all possible events associated with this sample space.
   (b) Show that if we are given the probabilities $\mathbb{P}(\{\omega_1\}), \mathbb{P}(\{\omega_2\})$, then we may deduce the probabilities of all other events using the basic axioms.
   (c) Similarly, show that if we are instead given the probabilities $\mathbb{P}(\{\omega_1, \omega_2\}), \mathbb{P}(\{\omega_1, \omega_3\})$, then we may deduce the probabilities of all other events using the basic axioms.

6. In an experiment, two weighted coins are flipped. It is given that both coins are weighted in the same way. After many repetitions, it is found that 12% of the trials result in both coins showing heads, and 15% of the trials result in both coins showing tails. Determine the probability that if one of the coins is flipped, it shows heads.

7. (a) From the axioms for a probability measure, prove that for any two events A, B (not necessarily exclusive), we have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. Generalise this formula to three events A, B, C.
   (b) A card is drawn randomly from a standard pack. Using the generalised formula in part (a), determine the probability that the card either shows a prime number, is a spade, or is red.

[1] These axioms are called the *Kolmogorov axioms*.

Conditional probability

8. Suppose that S is a (discrete) sample space, $\mathcal{F}$ is the set of all events, and $\mathbb{P} : \mathcal{F} \to [0,1]$ is a probability measure.
   (a) Define the *conditional probability* $\mathbb{P}(B|A)$ of an event B given an event A.
   (b) Show that the conditional probability function $\mathbb{P}(\cdot|A) : \mathcal{F} \to [0,1]$ satisfies the three basic axioms for a probability measure.
   (c) State the definition for two events A, B being *independent* under a probability measure, and explain why this definition makes sense using the definition of conditional probability.
9. A box of 100 gaskets contains ten gaskets with type-A defects only, five with type-B defects only, and two with both types of defect. Given that a gasket drawn at random has a type-A defect, what is the probability that it also has a type-B defect?
10. Your supervisor has two children, who are either boys or girls. Assuming equal probability of either gender, determine: (a) the probability that at least one child is a boy, given that at least one is a girl; (b) the probability that at least one child is a boy, given that the *younger* child is a girl.

Bayes' theorem

11. State and prove *Bayes' theorem*. Give an interpretation of each of the terms that arise.
12. You randomly choose a biscuit from one of two seemingly identical jars. Jar A has 10 chocolate biscuits and 30 plain; jar B has 20 chocolate and 20 plain biscuits. Unfortunately, you choose a plain biscuit. What is the probability that you chose from jar A?
13. Suppose that a disease affects one person in a thousand, and that a medical test for the disease accurately classifies 99% of all cases. What is the probability that, in a random screening exercise, a person who tests positively for the disease actually has the disease?

Combinatorics

14. (a) How many ways are there to order n distinct objects?
   (b) How many ways are there to choose r distinct objects from a set of n distinct objects, if ordering matters?
   (c) How many ways are there to choose r distinct objects from a set of n distinct objects, if ordering is unimportant?
   (d) How many ways are there to divide n identical objects into r groups?
15. Prove that in a group of n people, the probability of two sharing the same birthday is given by:
   $$1 - \frac{{}^{365}P_n}{365^n}.$$
   Using a calculator, determine the value of n for which this first exceeds $\frac{1}{2}$.
16. In one of the National Lottery games, six balls are drawn at random from 49 balls, numbered from one to 49. You pick six different numbers.
   (a) What is the probability that your six numbers match those drawn?
   (b) What is the probability that exactly r of the numbers you choose match those drawn?
   (c) What is the probability that five numbers of those you choose match those drawn and that your sixth number matches a 'bonus ball' drawn from those remaining after the first six balls are drawn?
17. Suppose that n distinguishable particles are placed randomly into N different states. A particular configuration of this system is such that there are n_s particles in state s, where $1 \leq s \leq N$. If the ordering of particles in any particular state does not matter, show that the number of ways of realising a particular configurations is:
   $$\frac{n!}{n_1!n_2!...n_N!}.$$

Discrete random variables

18. Let S be a sample space. A *discrete random variable* is a function $X : S \to T$ where T is a discrete set (often the set of integers). We define the *probability* of X taking a value in the set $T' \subseteq T$ to be:

$$\mathbb{P}(X \in T') := \mathbb{P}\left(\{\omega \in S : X(\omega) \in T'\}\right).$$

The *probability mass function* $f_X : \mathbb{Z} \to [0,1]$ of a discrete random variable is then defined to be:

$$f_X(x) = \mathbb{P}(X = x).$$

Using this formal definition, compute the probability mass functions of the following random variables:

(a) X_1, defined to be the number of heads observed when three fair coins are flipped;

(b) X_2, defined to be the number of primes observed when three fair six-sided dice are rolled.

19. Define the *expectation* $\mathbb{E}[X]$ and *variance* $\mathrm{Var}[X]$ of a discrete random variable X. Prove that for any two discrete random variables X, Y, and any constants a, b, we have $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

20. Let X be the result of a roll of a biased die, which displays one with probability $p/2$, two, three, four or five with probability p, and six with probability $2p$. Compute $\mathbb{E}[X]$ and $\mathrm{Var}[X]$.

21. After a late night out in Cambridge, you attempt to open the door to your house with the three keys in your pocket, only one of which is the correct door key. Assuming that you are equally likely to select any of the keys in your pocket, compute the probability mass functions, the expectations, and variances of:

(a) the random variable X, which is the number of attempts required to open the door if once you try a key from your pocket, you discard it on the ground if it doesn't work;

(b) the random variable Y, which is the number of attempts required to open the door if once you try a key from your pocket, you place it back in your pocket again.

The binomial and Poisson distributions

22. (a) Explain what it means to say that a random variable X has a *binomial distribution* $B(n,p)$. Derive a formula for the expectation and variance of a binomial distribution.

(b) An opaque bag contains 10 green counters and 20 red. One counter is selected at random and then replaced: green scores one and red scores zero. Five draws are made. If X is the total score, determine its expectation and variance.

23. (a) Explain what it means to say that a random variable X has a *Poisson distribution* $\mathrm{Pois}(\lambda)$. Derive a formula for the expectation and variance of a binomial distribution.

(b) Show that the limit of a binomial distribution $B(n,p)$ as $n \to \infty$, $p \to 0$ with $np = \lambda$ fixed, is a Poisson distribution $\mathrm{Pois}(\lambda)$.

Continuous random variables

24. (a) Given the *probability density function* $f_X(x)$ of a continuous random variable X, define its *expectation* $\mathbb{E}[X]$ and *variance* $\mathrm{Var}[X]$.

(b) If X is a continuous random variable with probability density function:

$$f_X(x) = \begin{cases} 10dx^2, & 0 \le x \le \frac{3}{5}, \\ 9d(1-x), & \frac{3}{5} \le x \le 1, \\ 0, & \text{otherwise,} \end{cases}$$

compute the most likely value of X, the expectation $\mathbb{E}[X]$, and the variance $\mathrm{Var}[X]$.

25. The lifetime, t, of a single bulb in a traffic signal is a random variable with density $f_T(t) = 1$ if $1 \leq t \leq 2$, and $f_T(t) = 0$ otherwise, where t is measured in years.
    (a) What is the probability as a function of y that a bulb fails in less than y years?
    (b) If the traffic signal contains three bulbs, and they fail independently, what is the probability as a function of z that none of the bulbs have to be replaced in z years?
    (c) What is the probability that the sum of the lifetimes of the three bulbs in a traffic signal exceeds 4 years?

26. Explain what it means to say that a random variable X is *normally distributed*, written $X \sim N(\mu, \sigma^2)$. Derive a formula for the expectation and variance of a normal distribution.

Past paper questions

Only attempt these questions before the supervision if you have lots of spare time. Sometimes, we might do them together in the supervision; otherwise, you can use these questions as revision material in the holidays.

Paper 1, Question 16, 2019 (20 marks)
A box contains 2 blue and 3 non-blue but otherwise identical balls. An experiment consists of three consecutive events: drawing a ball from the box, returning or not returning it back to the box, and drawing a second ball from the box. For example, and experiment might consist of (i) the event B_1 of drawing a blue ball in a first draw, (ii) the event R of returning the ball to the box and (iii) the event $\overline{B}_2$ of drawing a non-blue ball in a second draw. The probability of event R is $P(R) = r$.

(a) Find the sample space of this experiment and the probabilities of all possible outcomes using notations such as $P(B_1 \cap \overline{R} \cap B_2)$ for the probability of outcome $B_1 \cap \overline{R} \cap B_2$.

(b) Find the probability $P(B_2)$ that the second ball is blue if: (i) $r = 0$, (ii) $r = 1$, (iii) r is arbitrary in the range $0 \leq r \leq 1$.

(c) Find (i) $P(B_1 \cap B_2)$, (ii) $P(R|B_1 \cap B_2)$.

(d) Consider the general case of a box containing $N_B > 1$ of blue and $N - N_B > 1$ of non-blue but otherwise identical balls. For the experiment described above, find $P(R|\overline{B}_1 \cap B_2)$ in this general case. By sketching this probability as a function of r for fixed N and N_B, show that $P(R|\overline{B}_1 \cap B_2) \leq r$.

Paper 2, Question 14, 2021 (20 marks)
A coffee machine is supposed to serve drinks containing 300 ml of liquid. Coffee machines are verified by a test consisting of a single trial (producing a single cup of coffee), and the machine passes if it serves between 290 and 310 ml of liquid. A particular coffee machine makes random and independent servings in the range between 270 and 330 ml. The volume of liquid in each serving can be considered as a continuous random variable, v.

(a) Suppose the probability density function for v is:

$$P(v) = \frac{\pi}{120} \cos\left(\frac{\pi(v - 300\,\text{ml})}{60\,\text{ml}}\right) (\text{ml})^{-1}$$

between 270 ml and 330 ml (and zero otherwise).

   (i) Sketch the graph of the probability density function for v.
   (ii) Determine the probability that the machine passes the verification test.
   (iii) Prove that if $0 < r < 1$, then $\displaystyle\sum_{j=1}^{\infty} j r^{j-1} = (1 - r)^{-2}$.
   (iv) If the verification is repeated, used the formula in (iii) to calculate the expected number of tests in order for the fault to be discovered.

(b) Two coffee machines are made with the machine. Calculate the probability that the total volume of liquid is at least 630 ml.