Part IA: Mathematics for Natural Sciences B
Examples Sheet 8: Probability spaces, conditional probability, and combinatorics

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. Throughout, a bar over the top of an event denotes its complement within a sample space.

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 roll, 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) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

4. Suppose that S is a sample space and A is an event. Simplify the expressions:

   (a) $A \cap S$, (b) $(A \cap \overline{A}) \cup (A \cup \overline{A}) \cup \overline{A}$.

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 *Kolmogorov axioms* that a probability measure $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ must satisfy?

   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. (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.

Conditional probability

7. Suppose that S is a (discrete) sample space, $\mathcal{F}$ is the set of all events, and $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is a probability measure.
   (a) Define the *conditional probability* $\mathbb{P}(B|A)$ of an event B given an event A.
   (b) Show that, for a fixed A, the conditional probability function $\mathbb{P}(\cdot|A) : \mathcal{F} \to \mathbb{R}$ 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.
8. 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?
9. 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

10. State and prove *Bayes' theorem*. Give an interpretation of each of the terms that arise.
11. 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?
12. Three standard six-sided dice are tossed once onto a table. Let T be the event that two of the dice show the same value, and one is different. Let F be the event that the sum of the values shown by the dice is five or less. Calculate the probability $\mathbb{P}(F|T)$.
13. (**The base rate fallacy**) 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 order r objects from a set of n distinct objects?
   (c) How many ways are there to choose a subset of r objects from a set of n distinct objects?
   (d) How many ways are there to arrange n identical objects into r groups?
15. Prove the following property of the binomial coefficients:
$$\binom{n+1}{r+1} = \binom{n}{r+1} + \binom{n}{r},$$
   using: (a) the expression for the binomial coefficients in terms of factorials; (b) the combinatorial interpretation of the binomial coefficients in terms of combinations. Explain how this property of binomial coefficients relates to *Pascal's triangle*.
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:

$$\frac{n!}{n_1!n_2!...n_N!}.$$

18. Letters A, B, C, D, E, and F are written in a random order, but without repetition, into places $1, 2, 3, 4, 5$ and 6. Explaining your reasoning in each case, how many distinct orderings:

    (a) exist in total?

    (b) have F in the sixth place?

    (c) have E or F in the sixth place?

    (d) have E in the fifth place, and F in the sixth place?

    (e) have E in the fifth place, or F in the sixth place?

    (f) have E in the fifth place, or F in the sixth place, but *not both*?

    Now, the letters A, B, C, D, E and F are instead partitioned into two bins, where order does not matter in a given bin. We say that a partition is of *type* $[a, b]$, if a letters are placed into the first bin, and b letters are placed into the second bin.

    (g) How many partitions of type $[4, 2]$ are there?

    (h) Assuming that from all the possible orderings given enumerated in part (a), the letters in the first four places are placed into the first bin, and the letters in the final two places are placed into the second bin, how many times do A, B, C, D end up in the first bin overall?

    (i) Calculate the product of your answers to the two previous parts, and explain the value you obtain.

    (j) Repeat the calculation of parts (g)-(i) for each of the possible types of partitions.

Miscellaneous probability space problems

[*This section contains various longer questions, based on past tripos questions. Question 21, related to sampling with and without replacement, has come up a lot - it is well worth having a proper go at that one!*]

19. A box contains $N_B \geq 2$ blue balls and $N - N_B \geq 2$ non-blue balls. An experiment consists of three consecutive stages: drawing a ball from a box, returning it or not returning it, then drawing a second ball from the box. The event B_i represents a blue ball being drawn on the ith draw, for $i = 1, 2$. The event R represents returning a ball on the second stage of the experiment. The probability of event R is $\mathbb{P}(R) = r$.

    (a) Write down the sample space of the experiment, and find the probabilities of all of the possible outcomes.

    (b) Hence, find:

        (i) $\mathbb{P}(B_2)$;

        (ii) $\mathbb{P}(B_1 \cap B_2)$;

        (iii) $\mathbb{P}(R|B_1 \cap B_2)$.

    (c) By sketching the graph of $\mathbb{P}(R|\overline{B}_1 \cap B_2)$ as a function of r, show that $\mathbb{P}(R|\overline{B}_1 \cap B_2) \leq r$.

20. A factory produces good bananas with probability p and bad bananas with probability $1-p$. The bananas are placed on a conveyor belt and inspected by n different workers sequentially. Worker k notices a good banana with probability g_k, and removes it from the conveyor belt in this instance. Worker k notices a bad banana with probability b_k, and also removes it from the conveyor belt in this instance. Assume that $0 < p < 1$ and that $g_k > 0, b_k > 0$ for all $k = 1, ..., n$.

    An experiment is conducted where a banana, which may be good or bad, is placed on the conveyor belt for inspection.

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

    (b) Let G be the event that a good banana is produced, and let X_k be the event that the banana is removed from the conveyor belt by worker k. Find:

    (i) $\mathbb{P}(G \cap X_1)$;
    (ii) $\mathbb{P}(X_1)$;
    (iii) $\mathbb{P}(\overline{G}|X_2)$;
    (iv) $\mathbb{P}(G|X_2 \cup X_3 \cup ... \cup X_n)$;
    (v) $\mathbb{P}(X_{k+1} \cup X_{k+2} \cup ... \cup X_n)$;
    (vi) $\mathbb{P}(G|X_1 \cup X_2 \cup ... \cup X_k)$.

    (c) If $b_1 = b_2 = ... = b_n$, and $n = 98$, find the minimal value of p for which $\mathbb{P}(G|X_1) = 0.99$.

21. (**Sampling with replacement**) A bag is filled with N white balls and M black balls. Balls are drawn from the bag sequentially without replacement. Let W_i denote the event that the ith ball drawn is white, and let B_i denote the event that the ith ball drawn is black.

    (a) Find $\mathbb{P}(W_1)$, $\mathbb{P}(W_2)$, and $\mathbb{P}(W_3)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_i)$.

    (b) What does the event $W_i \cap W_j$ represent? Find $\mathbb{P}(W_1 \cap W_2)$, $\mathbb{P}(W_1 \cap W_3)$ and $\mathbb{P}(W_2 \cap W_3)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_i \cap W_j)$.

    (c) What does the event $W_i \cup W_j$ represent? Find $\mathbb{P}(W_1 \cup W_2)$, $\mathbb{P}(W_1 \cup W_3)$ and $\mathbb{P}(W_2 \cup W_3)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_i \cup W_j)$.

    (d) Using Bayes' theorem, compute $\mathbb{P}(W_2|W_1)$ and $\mathbb{P}(W_3|W_2)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_{i+1}|W_i)$.

    (e) Show that the probability of obtaining exactly $n \leq N$ white balls in a total of x draws is given by:

$$\frac{\binom{N}{n}\binom{M}{x-n}}{\binom{N+M}{x}}.$$