1 Some challenges in combinatorics and probability

1.1 Introduction

We make the following definition for an integer $n \geq 1$:

$$n! = n(n-1)(n-2)\cdots 2 \cdot 1, \quad \text{with } 0!=1.$$

The ! operation is called *factorial*. The number $n!$ gives the number of ways of arranging n distinct objects in a line. An equivalent definition is the number of ways you can choose n distinct objects from a set (without replacement), and where the order of selection matters. The following is called a *binomial coefficient*:

$$\binom{n}{r} := \frac{n!}{(n-r)!r!}, \quad \text{with } 0 \leq r \leq n.$$

Question: what does this count?

The following is called a *multinomial coefficient*:

$$\binom{n}{n_1!, n_2!, \ldots n_k!} := \frac{n!}{n_1! n_2! \cdots n_k!}, \quad \text{with } n_1 + n_2 + \cdots + n_k = n.$$

Question: what does this count?

I will spend about 10 minutes outlining the above. You are then welcome to try all other sections in any order as they are accessible to all. Generally the sections get more challenging as we go along, and I'd advise most of you to start with Section 1.2: as I'll focus on these questions in any early discussions. Myself and helpers can give hints for those doing questions in the later sections. I won't have time to cover everything. I would like to finish on the Monty Hall Problem in Section 1.6, so feel free to jump ahead to that section too.

1.2 Some elementary counting and probability problems

Question: Write down some values for $n!$, e.g. for $n \leq 10$. Write down

$$\binom{100}{97}, \ \binom{10}{3,3,4}.$$

Question: Find the number of distinct arrangements of the following sequence of letters:

ABCDE, AAABBBB, ABBCCC.

Question: Let $S = \{(a,b) : 1 \leq a, b \leq 10\}$, where (a,b) is a pair of integers. Let $\mathcal{A} = \{(1,b) : 1 \leq b \leq 10\}$, and $\mathcal{B} = \{(a,1) : 1 \leq a \leq 10\}$. Write down the number of elements in the following sets:

$$\mathcal{A} \cap \mathcal{B},\ \mathcal{A} \cup \mathcal{B},\ \mathcal{A}^c,\ \mathcal{B}^c,\ \mathcal{A}^c \cap \mathcal{B}^c,\ \mathcal{A}^c \cup \mathcal{B}^c,\ \mathcal{A} \cap \mathcal{B}^c,\ \mathcal{A} \cup \mathcal{B}^c.$$

Here A^c means the complement of A, i.e. the set of elements in S not in A.

Question: Consider the set $S = \{1,2,3,4,5\}$. Suppose I choose 3 numbers from S without replacement, and note them down. Suppose a second person tries to guess what I chose by listing 3 numbers from S. i) How many possible combinations of numbers can the second person choose that match at least two of my chosen numbers? ii) If each combination of 3 numbers is equally likely to occur, what is the probability that the second person will choose 3 numbers that match at least two of my chosen numbers? iii) How about at least one of my chosen numbers?

Question: A family has three children, and each child is found to be born on the same day of the year. All are different ages, so none are twins nor triplets. In a newspaper article it reported that there is approximately a 1 in 50 million chance that this event will occur. Criticise this argument. (You may assume that children are equally likely to be born on any day of the year).

Question: A family has two children, and each child is found to be born at the same time of day (to nearest minute). On the radio news it is reported that there is approximately a 1 in 2 million chance that this event will occur. Criticise this argument. (You may assume that children are equally likely to be born at any time of day).

Question: How many people do you need in a room to ensure that the probability of at least two people sharing a birthday is at least 1/2? (You may assume that each person is equally likely to be born on any day of the year. We'll also ignore leap years).

Question: Consider the set $S = \{1, 2, 3, \ldots, 10\}$. Suppose I choose 4 numbers from S without replacement, and note them down. Suppose a second person tries to guess what I chose by listing 4 numbers from S. i) How many possible combinations of numbers can the second person choose that match at least one of my chosen numbers? ii) If each combination of 4 numbers is equally likely to occur, what is the probability that the second person will choose 4 numbers that match exactly three of my chosen numbers?

Question: Consider the sequence of letters $YOUME$. Let $\mathcal{A}$ denote the set of sequences containing YOU, and let $\mathcal{B}$ denote the set of sequences containing ME. If $|\cdot|$ denotes the number of elements in a set, find: $|\mathcal{A} \cap \mathcal{B}|$, $|\mathcal{A}|$, $|\mathcal{B}|$, $|\mathcal{A} \cup \mathcal{B}|$ and $|(\mathcal{A} \cup \mathcal{B})^c|$.

1.3 Further counting problems

Question: Without using any algebra explain using counting arguments why:

$$\binom{n}{r} = \binom{n}{n-r}.$$

[Hint: for each choice of r objects from n, consider the objects that are left behind.]

Question: Without using any algebra explain using counting arguments why:

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}.$$

[Hint: Label one object as being special. When you choose r objects from n, consider how many choices include the special object, and how many choices exclude the special object.] Optional: confirm your result by the algebra of factorials.

By extending the argument above deduce that

$$\binom{n}{r} = \binom{n-2}{r} + 2\binom{n-2}{r-1} + \binom{n-2}{r-2}.$$

Question: There are three distinct boxes (labelled Box 1, Box 2 and Box 3), and 5 identical red balls. How many *distinct* ways are there of putting the 5 balls in the three boxes. Empty boxes are allowed.

[Hint: 3 balls in Box 1, 0 balls in Box 2, and 2 balls in Box 3 can be represented by the string 0001100. 2 balls in Box 1, 2 balls in Box 2, and 1 ball in Box 3 can be represented by the string 0010010. Thus the 1s can be thought of as dividers between the boxes, and the 0s are the balls in each box. How many strings are there?]

Question: Suppose there are r distinct boxes and n identical balls. What is the general formula for the number of distinct arrangements of putting the balls into the boxes? (Empty boxes allowed).

Question: Suppose there are r distinct boxes and n identical balls. If empty boxes *are not allowed*, show that the general formula for the number of distinct arrangements of putting the balls into the boxes is given by:

$$\binom{n-1}{r-1}.$$

[Hint: Given a sequence of n zeros, how many ways are there of inserting $(r-1)$ 1s to segregate the zeros into sub-sequences containing at least one zero?]

Question: How many distinct terms are there in the expansion of $(w+x+y+z)^5$? [Hint: the above analysis is useful].

1.4 Steiner systems

This is about Steiner systems. A Steiner system with parameters r, k, and n, written $S(r,k,n)$, is an n-element set S together with a set of k-element subsets of S (called blocks) with the property that each r-element subset of S is contained in exactly one block. A problem in combinatorics is to construct such a list of blocks given these parameters.

Example: given the subset $S = \{1,2,3,4\}$ we want to consider the minimum number of blocks $A_1, A_2, \ldots A_k$ each with 3-elements, such that any 2-element subset of S can be found in exactly one of the A_i. In this example, there are 6 subsets of S of size 2. Each A_k, being of size 3 will contain 3 subsets of size 2. e.g. $A_1 = \{1,2,3\}$ contains the subsets $\{1,2\},\{1,3\}$ and $\{2,3\}$. We see that within

$$A_1 = \{1,2,3\},\ A_2 = \{1,2,4\},\ A_3 = \{2,3,4\},$$

we can find all subsets of size 2. However in this list the sets $\{1,2\},\{2,3\}$ and $\{2,4\}$ can be found twice, so the example above is not a Steiner $S(2,3,4)$ system.

Question: Is it possible to refine the list above so that every subset of size 2 appears only once? Hence is $S(2,3,4)$ a Steiner system?

Question: Consider the set $S = \{1,2,3,4,5\}$. Is $S(2,3,5)$ a Steiner system? Hint: first work out how many subsets of S there are of size 2, noting that each 3-element block A_i contains 3 subsets of size 2. Harder: what is the minimum number of sets A_i required such that all subsets of S of size 2 can be found in at least one of the A_i? (I've not worked this out!)

Question: If $S(r,k,n)$ is to form a Steiner system show that the number B of blocks required is given by:

$$B = \frac{\binom{n}{r}}{\binom{k}{r}}.$$

Notice that there exist examples where B is an integer but $S(r,k,n)$ is not a Steiner system. Clearly if $B \notin \mathbb{N}$ then $S(r,k,n)$ will not form a Steiner system.

Question: Show that $S(2,3,7)$ and $S(2,3,9)$ are Steiner systems. i.e. construct the blocks A_i so that every 2-subset of S appears in one and only one of the A_i. [A hint will be given in class for those that try this. First work out how many sets A_i will be required.]

Open Problem. Given the set $S = \{1, 2, \ldots, 49\}$, construct a list A_1, $A_2 \ldots, A_k$ of subsets of size 6, such that all subsets of size 3 are represented at least once. What is the minimum number k required? This problem is connected to the minimum number of tickets needed to ensure a win in the UK national lottery. (However the minimal list will not lead to a positive expected profit!). We remark that $S(3, 6, 48)$ is a Steiner system.

1.5 Coin tossing sequences (and probability)

A coin tossing sequence is a list of the form $HHTTH\ldots$, where H =Heads and T =Tails. i.e. it is a list of outcomes of consecutive coin tossings. We will assume coins are fair so that Heads is equally as likely as Tails (and each coin toss is independent of another).

Question: How many coin tossing sequences are there of length n? If Tails followed by Tails is not allowed how many coin tossing sequences are there of length n in this case? [Hint: First work out allowed sequences of length up to five and then make a guess for the general case.] Harder: prove your guess.

Question: A single coin is tossed over and over again and the sequence of Heads and Tails recorded. Player 1 chooses the sequence HH, while Player 2 chooses the sequence TH. A player wins if their sequence appears first. Who is more likely to win (or are they equally likely)?

Question: Repeat the above question if Player 1 chooses HHH, while Player 2 chooses THH. What about Player 1 choosing HHHH versus Player 2 choosing THHH? Harder: consider the set of coin tossing sequences of length three:

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

If Player 1 is first to choose a sequence from S, can Player 2 always choose a sequence from S that's more likely to appear before Player 1's sequence?

1.6 The Monty Hall Problem

Given time (perhaps in the last 10 minutes of the session) I would like to discuss this classical problem, based on a TV game show. The game goes like this:

You are presented with three doors (numbered 1,2 and 3). Hidden behind two of the doors are goats, and hidden behind a remaining door is a car. You don't know what is behind a particular door and you want to win the car. The game show host invites you to pick a door. The host goes on to explain that once you pick a door, he (the host) will open another door to reveal a goat. The host will not reveal what is behind your door, and the host knows where the car is hidden. The host will invite you to change your mind.

You pick door number 1, the host then opens door number 2 to reveal a goat. The host then invites you to change your mind from opening door 1 to opening door number 3. Should you switch choice?