II Probability: lecture 2

II.1 Basic Counting Principles

Suppose you are the person responsible for assigning license plate numbers for cars in the state of New York. The plate looks like the figure below: it has 3 letters and 4 digits. You are interested in finding the maximum number of license plates that can be issued in NY. To calculate this, you have to count all possible outcomes of obtaining 3 letters followed by 4 digits.

We will assume that we are working with the english alphabet, which has 26 letters, and we have ten distinct digits, 0 through 9. Therefore, to choose the first letter on the plate, we have 26 options. Similarly, we have 26 options for the second letter, and 26 for the third. As for the four digits, we have 10 options for each digit. This gives a total of

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 \times 10 = 175,760,000 \tag{1}$$

possible license plate numbers. Note that we have not imposed any restrictions on the repetition of letters and numbers in this calculation. If repetition were prohibited, there will still be 26 choices for the first letter on the plate. However, once that letter is chosen, it cannot be used again, thus leaving only 25 choices for the second letter, and 24 for the third. A similar argument can be applied to the digits on the plate, which leads to a total of

$$26 \times 25 \times 24 \times 10 \times 9 \times 8 \times 7 = 78,624,000 \tag{2}$$

possible license plate numbers.

II.2 Permutations

A *permutation* or an *ordered list* is the arrangement of the elements of a given set into a specific order. Let S be a set which has the elements a, b and c. One possible permutation of S is $\{c, a, b\}$, another one is $\{a, b, c\}$. We are interested in finding the total number of ordered lists of the set S. Here, order is very important, and repetition of elements is prohibited. For example, if we choose c to be the first element in the list, then the second element must be either a or b. When choosing the first element in a permutation of S, we have three different possibilities, a, b or c. Once the first element is chosen, we are left with two options for the second element of the permutation, and when that is selected, one option remains for the last entry. Therefore, the total number of permutations of a set made of three elements is

$$3 \times 2 \times 1 = 6. \tag{3}$$

If the set had five elements in it, then the total number of permutations would be $5 \times 4 \times 3 \times 2 \times 1 = 120$. For an arbitrary number of elements n there are

$$n \times (n-1) \times (n-2) \times (n-3) \times \cdots \times 3 \times 2 \times 1 = n! \tag{4}$$

permutations, where the notation "n!" reads "n-factorial".

Example: Suppose you go with 9 of your friends to watch a play. You have booked the entire fifth row in the theatre, which has exactly 10 seats. In how many possible ways can you and your friends occupy the seats of the fifth row?

The answer is simple: whoever shows up first has 10 seats to choose from, the second person to arrive has 9 seats, and so on. The total number of ways that the 10 of you can be seated is $10! = 3,628,800$.

In summary, given a set of n elements, you can create a total of $n!$ ordered lists.

II.3 Total number of subsets of a set

You have a set which has a total of n elements, and you would like to find the total number of all possible subsets of this set. For example, let the set $S = \{1, 2, 3, 4, 5, 6\}$, then $\{1\}$, $\{2, 3\}$, $\{1, 5, 6\}$, and $\{\}$ (this is just an empty set) are all subsets of S. In order to find the total number of subsets, we need to look at each element of S and decide whether it is going to be a part of the subset or not. Therefore, each element of S has two options: it is either *in* the subset, or it is *out* of the subset. For the given set $S = \{1, 2, 3, 4, 5, 6\}$, the first element has two options (either in or out of the subset), and so do the remaining five elements. Therefore, the total number of subsets of S is

$$2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^6 = 64. \tag{5}$$

In general, for a set of n elements, the total number of subsets 2^n.

II.4 Combinations

Now that we have seen how many ordered lists, and how many different subsets a given set has, we can proceed to find how many ordered lists of k elements can be created from a set of n elements, where $k \leq n$. We will do this for a given k and then we will generalize.
Suppose you are in a classroom with $n = 70$ students, and the professor decides to create a class committee made of $k = 5$ students. How many possible committees can the professor come up with? Think of it as follows: there are five chairs which the professor would like to assign students to. Think of the chairs as numbered 1 through 5, where the person assigned

to chair 1 will be the president of the committee, chair 2 the vice-president, and so on. The professor has 70 students to choose from, and a student can be selected for any of the five available seats.
There are two ways to calculate the total number of ordered lists of 5 students from a total of 70:

1. The first way is done by finding the total number of ways to fill five chairs (in order) from the 70 available students.

2. The second way is done by first choosing 5 out of 70 students, and then ordering those chosen 5.

The first calculation is easier to perform, so we proceed with that:

- There are 70 choices to fill chair 1.
- There are 69 choices to fill chair 2.
- There are 68 choices to fill chair 3.
- There are 67 choices to fill chair 4.
- There are 66 choices to fill chair 5.

Thus, the total number of ways to fill those five chairs, i.e., the total number of ordered lists of 5 students from a total of 70 students is

$$70 \times 69 \times 68 \times 67 \times 66 = 1,452,361,680 = \frac{70!}{(70-5)!}. \tag{6}$$

The second method can be done by first calculating all the possible ways in which we can choose 5 out of 70 students. We do not know how to calculate this yet, but we will denote it by $\binom{70}{5}$, pronounced "70 choose 5". Now that we have chosen the 5 students to fill the chairs, the only thing left to do is deciding who is president, who is vice-president, etc., in other words creating an ordered list. As we have seen in other examples, the answer will be the total number of permutations of 5 chairs, which is 5!. Thus, the total number of ordered lists of 5 students from a total of 70 students is

$$\binom{70}{5} \times 5!. \tag{7}$$

This number must equal the one calculated from the first method in equation (6), i.e.,

$$\binom{70}{5} \times 5! = \frac{70!}{(70-5)!}. \tag{8}$$

Therefore, we can now define how to choose 5 students out of a total of 70 as

$$\binom{70}{5} = \frac{70!}{5!(70-5)!}, \tag{9}$$

and this is called a combination. In general (for any $k \leq n$, where k and n are integers), there are

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{10}$$

ways to choose k elements out of a set of n elements, when we care only about which items we selected but not about the order in which we picked them out. These symbols $\binom{n}{k}$ are also called "choose symbols" or "binomial coefficients". By convention $0! = 1$. Thus $\binom{n}{0}$ has the meaning $\frac{n!}{n!0!} = 1$.

These coefficients lead to many nice properties, of which we will only point out one. Recall that the total number of all possible subsets of a set with n elements is 2^n, and the total number of subsets made of k elements is $\binom{n}{k}$. If k takes all values from 0 to n, then the elements $\binom{n}{k}$ give the total number of subsets of size 0, 1, 2, ..., n. In other words $\binom{n}{k}$ give the total number of **all** subsets, i.e.

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-2} + \binom{n}{n-1} + \binom{n}{n} = 2^n. \tag{11}$$

II.5 Birthday coincidences

We are interested in finding the minimum amount of people it takes to find a pair of coinciding birthdays with probability greater than 0.5. Assume 1 year = 365 days, and no leap year birthdays allowed.

- If you meet a stranger, what is the probability that they have the same birthday as you? The answer is 1/365.
- If two people meet, what is the probability that they have the same birthday? The answer is also 1/365, because you want to compare the first person's birthday to that of the second person, and since we know that the second person is born on one specific day in the year, then the chance that the first person's birthday matches is 1/365.
- Same for three people: what is the probability that at least two have same birthday? This one is not as straight forward, because now you have to do pairwise comparisons, i.e. compare first person with second, second with third, and first with third. The easier calculation to perform would be to find the probability that all three birthdays are different. To do this, you want to count all the possible birthdays each of the three people has. The first person can be born on any of the 365 days. The second person can be born any day expect the day the first person is born, therefore the second person can be born on any of 364 days. Similarly, the third person can be born on any of 363 days. Thus, the probability that all three people have different birthdays is

$$P(\text{all three have different birthdays}) = \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365}.$$

Then, the probability that at least two people have the same birthday is

$$\begin{aligned} P(\text{at least two have the same birthday}) &= 1 - P(\text{all three have different birthdays}) \\ &= 1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \\ &= 0.0082 \text{ or } 0.82\%. \end{aligned}$$

- Four people: probability that at least two have same birthday?

$$\begin{aligned} P(\text{at least two have the same birthday}) &= 1 - P(\text{all four have different birthdays}) \\ &= 1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \\ &= 0.016 \text{ or } 1.6\%. \end{aligned}$$

- How many people would you guess that you need for this probability to reach 0.5? In general, if we have n people we find that

$$\begin{aligned} P(\text{at least two have the same birthday}) &= 1 - P(\text{all } n \text{ have different birthdays}) \\ &= 1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \cdots \times \frac{365-n+1}{365} \\ &= 1 - \frac{365!}{365^n(365-n)!}. \end{aligned} \tag{12}$$

If we try several values for n in equation (12), we find the following:

· $n = 5 \implies P(\text{at least two have the same birthday}) = 0.027$

· $n = 10 \implies P(\text{at least two have the same birthday}) = 0.117$

· $n = 15 \implies P(\text{at least two have the same birthday}) = 0.253$

· $n = 20 \implies P(\text{at least two have the same birthday}) = 0.411$

· $n = 22 \implies P(\text{at least two have the same birthday}) = 0.476$

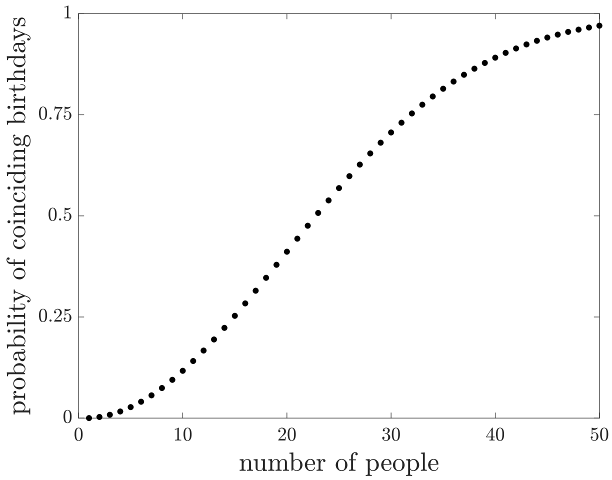
· $n = 23 \implies P(\text{at least two have the same birthday}) = 0.507$

Thus, if we have 23 people, there is more than 50% chance to find coinciding birthdays. The figure below shows that if there are 50 people, the probability that we will find coinciding birthdays is close to 1.

II.6 Distributing cards

We have a deck of 52 cards, which we would like to distribute among 4 people. We are interested in calculating the probability that each person gets an ace.

To proceed with the solution, we note that to calculate the probability that each person gets an ace, we must first find the total number of ways we can distribute the 52 cards giving



one ace to each person. Once we have that number, we can divide it by the total number of ways we can distribute 52 cards among 4 players. Let's proceed with the easier part, that is finding the total number of ways we can distribute 52 cards.

We distribute by giving each player 13 cards, thus:

- We give 13 out of the total 52 cards to player 1. The number of ways to do this is $\binom{52}{13}$.
- We have 39 cards left, so we give 13 out of 39 cards to player 2. The number of ways to do this is $\binom{39}{13}$.
- We have 26 cards left, so we give 13 out of 26 cards to player 3. The number of ways to do this is $\binom{26}{13}$.
- We are left with 13 cards which we give to player 4. The number of ways to do this is $\binom{13}{13}$.

Thus, the total number of ways of distributing 52 cards, giving each player 13 of them, is

$$\binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13} = \frac{52!}{(13!)^4}. \tag{13}$$

The second part of the solution is to calculate in how many ways we can distribute 52 cards such that each player gets one ace. We first ask in how many ways we can distribute 4 aces to 4 players. This is simply a permutation, and the answer is 4!. We are now left with 48 cards, and we have to give each player 12 cards. Following the steps above, we proceed as follows:

- We give 12 out of the total 48 cards to player 1. The number of ways to do this is $\binom{48}{12}$.
- We have 36 cards left, so we give 12 out of 36 cards to player 2. The number of ways to do this is $\binom{36}{12}$.

- We have 24 cards left, so we give 12 out of 24 cards to player 3. The number of ways to do this is $\binom{24}{12}$.
- We are left with 12 cards which we give to player 4. The number of ways to do this is $\binom{12}{12}$.

Thus, the total number of ways of distributing 48 cards, giving each player 12 of them, is

$$\binom{48}{12}\binom{36}{12}\binom{24}{12}\binom{12}{12} = \frac{48!}{(12!)^4}. \tag{14}$$

This indicates that the total number of ways to distribute the 52 cards giving each player an ace is

$$4! \times \binom{48}{12}\binom{36}{12}\binom{24}{12}\binom{12}{12} = \frac{4! \times 48!}{(12!)^4}, \tag{15}$$

and therefore, the probability that each person gets an ace is

$$\frac{\frac{4! \times 48!}{(12!)^4}}{\frac{52!}{(13!)^4}}. \tag{16}$$