Chapter 5

Combinatorics

> *"I think you're begging the question," said Haydock, "and I can see looming ahead one of those terrible exercises in probability where six men have white hats and six men have black hats and you have to work it out by mathematics how likely it is that the hats will get mixed up and in what proportion. If you start thinking about things like that, you would go round the bend. Let me assure you of that!" (Agatha Christie, The Mirror Crackd)*

This chapter is dedicated to combinatorics, which refers broadly to different ways of counting objects.

Suppose for example that you have two slots to be filled, and for the first slot, there are n_1 choices, while there are n_2 choices for the second slot. How many ways are there to fill up both slots? Well, for the first slot, we have n_1 choices, now for each of these, we still have n_2 choices, for a total of $n_1 n_2$ choices.

Example 51. Suppose you have 3 choices for the main course, and 2 choices for the dessert. How many choices of menus do you have? Well, you can pick any of the 3 main courses, so 3 choices here. Next, for main course 1, you can choose 2 desserts, then for main course 2, you can choose 2 desserts, and finally for main course 3, you can still choose 2 desserts, which makes it a total of 6 menus.

More generally, if there are k slots, and n_1 choices for the 1st slot, n_2 choices for the second slot, until n_k choices for the kth slot, we get a total of $n_1 \cdot n_2 \cdots n_k$ choices.

Principle Of Counting

- There are two slots to be filled, there are n_1 *choices* for slot 1 and n_2 *choices* for slot 2.
  - E.g., you have 3 choices for the main course and 2 choices for dessert.
- The total number of unique choices to fill the slots is $n_1 n_2$
- In general: n_1, n_2, … n_k choices for *k-slots*
- $n_1 * n_2 * \ldots * n_k$ ways
  - (cardinality of the cartesian product of sets)

© image from http://key-boxes.com/page/2

Cardinality Of Power Set

- Consider a set A with n elements.
- Each of these n elements are *either in a subset of A or not:* 2 choices
  - Such a choice needs to be made for *each of the n elements*
- Thus $2*2*\ldots*2=2^n$ choices.
  - We saw another derivation using the Binomial theorem.

Now $n_1 \cdot n_2 \cdots n_k$ is also the cardinality of the cartesian product of k sets, where the set i has n_i elements.

Example 52. Suppose you have 3 choices for the main course, and 2 choices for the dessert. How many choices of menus do you have? An alternative way to view this question is to explicitly list all the choices:

main course 1, dessert 1	main course 1, dessert 2
main course 2, dessert 1	main course 2, dessert 2
main course 3, dessert 1	main course 3, dessert 2

This makes a total of 6 menus. You notice that when we list all the options, we get a cartesian product of two sets, the set { main course 1, main course 2, main course 3 }, and the set { dessert 1, dessert 2 }.

We recall that given a set A, its power set $P(A)$ is the set of all subsets of A. We already saw in the previous chapter that the cardinality of $P(A)$ is 2^n. Here is another way of proving it. Write $A = \{a_1, \dots, a_n\}$. Now list all subsets of A, and to each subset, associate a binary vector of length n, where every coefficient is either 0 or 1: the first coefficient is 1 if a_1 is in the subset, and 0 otherwise, similarly, the second coefficient is 1 if a_2 is in the subset, and 0 otherwise, and so on and so forth. Since every element is in a given subset or not, we do obtain all possible binary vectors of length n, and there are 2^n of them.

Example 53. Consider the set $A = \{1, 2\}$.

$$\begin{array}{c|l} \varnothing & 00 \\ \{1\} & 10 \\ \{2\} & 01 \\ A & 11 \end{array}$$

Now suppose that there are n elements, to be put in r slots. If elements can be repeated, we are in the scenario we have just seen, and there are n^r choices. Now if elements cannot be repeated, then, we have n choices for the first slot, $n-1$ choices for the second slot, and so on and so forth, until $n-(r-1)$ choices for the last slot. We thus get

$$n(n-1)(n-2)\cdots(n-r+1). \tag{5.1}$$

This is for example what happens when picking cards from a deck of cards, once the cards are not put back in the deck.

Filling r Slots With n Choices

- There are n elements, with which to fill *r* slots.
- When elements can be repeated:
  - Using the principle of counting: n*n*…*n = **n^r choices**
- When elements cannot be repeated:
  - n choices for first slot,
  - n-1 choices for second slot,…
  - n-(r-1) choices for last slot
  - In total: **n(n-1)(n-2)…(n-r+1)** choices

- **E.g., sequence of choice of cards from a deck of cards**

Permutation: P(n,r)

A permutation is an arrangement of all or part of a set of objects, *with* regard to the order of the arrangement.

Number of permutations of n objects taken r at a time:

$$\boxed{P(n,r) = n(n-1)(n-2)\ldots(n-r+1) = n!/(n-r)!}$$

where n!=n*(n-1)*(n-2)*…*2*1 (called n factorial).

If $r = n$, we notice that all the n elements are attributed to the n slots, which gives a *permutation* of the n elements. This also shows that the number of permutations of n elements is

$$n(n-1)(n-2)\cdots 2\cdot 1 = n!.$$

The above scenario, when there are n elements but only r slots, and elements cannot be repeated, is called *permutations of n objects taken r at a time*, that is an arrangement where ordering matters, and the number $P(n,r)$ of such permutations is

$$P(n,r) = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

Permutations can be defined also if some of the objects are repeated, that is, we still have n elements, but $n = k_1 + k_2 + \cdots + k_r$, that is, there are k_1 elements of type 1, k_2 elements of type 2, until k_r elements of type r. How many permutations do we have in this case? To count this, we can proceed as follows: place k_1 elements out of n places, then place k_2 elements in $n-k_1$ places, etc until you place k_r elements in the remaining places. This means that we have

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{k_r}{k_r} \tag{5.2}$$

where we recall that $\binom{n}{k}$ counts the number of ways of choosing k elements out of n.

We also call $\binom{n}{r}$ a *combination*, that is a say of selecting objectings without considering the order of the selection. We have that

$$\binom{n}{r} = C(n,r) = \frac{n!}{r!(n-r)!}.$$

Indeed, recall that when we had r slots and n objects, we have $P(n,r)$ ways of placing the objects, where

$$P(n,r) = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

Now we still have r slots, n objects to choose from, but this time, we do not care about the ordering, and there are $r!$ possible ordering for each combination:

$$r!C(n,r) = P(n,r).$$

Permutation

In general: The number of distinguishable permutations from a collection of objects, where first object appears (repeats) k_1 times, second object k_2 times, … for r distinct objects:

$$\boxed{n!/(k_1!\, k_2! \ldots k_r!)}$$

Combination: C(n,r) or $\binom{n}{r}$

A combination is a selection of all or part of a set of objects, *without* regard to the order in which objects are selected.

E.g. Team of 4 people from a group of 10

Number of combinations of n objects taken r at a time

$$\boxed{\binom{n}{r} = C(n,r) = n!/r!(n-r)!}$$

- There are r! possible orderings within each combination
- So r! C(n,r) = P(n,r) by definition of permutation

Example 54. From a committee of 8 people, in how many ways can you choose

- a chair and a vice-chair (one person cannot hold more than one position): it is $P(8,2) = \frac{8!}{6!} = 8 \cdot 7$. Indeed, once the chair is chosen (8 choices), we have 7 choices for the vice chair.
- a subcommittee of 2 people: it is $C(8,2) = \frac{8!}{2!6!} = 28$. Indeed, in this case, any 2 persons among the 8 people will do, irrespectively of the ordering. This means that we choose person 1 with person 2, person 1 with person 3, etc until person 1 with person 8 (7 possibilities),or person 2 with person 3, etc until person 2 with person 8 (6 possibilities), or person 3 with person 4, ..., person 3 with person 8 (5 possibilities), and by continuing the list, we get $7+6+5+4+3+2+1 = 28$.

This example also illustrates that $r!C(n,r) = P(n,r)$. Indeed, we know that $P(8,2)$ takes into account the ordering, therefore, if say person 1 is chair and person 2 is vice chair, it counts for 1 choice, while it counts for 2 choices for the subcommittee, since person 1 and person 2, and person 2 and person 1, represent the same subcommittee.

We now finish the computations of (5.2):

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{k_r}{k_r} = \frac{n!}{k_1!(n-k_1)!}\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdots\frac{(n-k_1-\cdots k_{r-1})!}{k_r!}$$

where we notice that we can cancel out numerator and denominator to finally obtain

$$\frac{n!}{k_1!k_2!\cdots k_r!}.$$

Some experiments may not have a deterministic outcome, e.g., tossing a coin, or throwing a dice, in which case, different techniques are needed.

Definition 34. The set of possible outcomes of random trial is called a sample space.

For example, if you toss a coin, the set of possible outcomes is { head, tail }. If you toss a coin twice, you may be interested in the number of heads, for which the set of possible outcomes is {0,1,2 }, or in the actually sequence of heads/tails, for which the set of possible outcomes is { head head, head tail, tail head, tail tail }, or in whether the two twosses matched, for an outcome which belongs to { yes, no }.

Example

From a committee of 8 people, in how many ways can you choose:

- a chair and vice-chair (one person cannot hold more than one position?)
- P(8,2)
- a subcommittee of 2 people?
- C(8,2)

Sample Space

- Some experiments may not have a deterministic outcome, e.g., tossing of a coin, throwing a dice.

The *set of possible outcomes of a random trial* is called the *sample space*.

- E.g, for coin toss, the sample space is {Head, Tail}
- Two coins tossed
  - Record the number of heads {0,1,2}
  - Record sequence of heads/tails {HH, HT, TH, TT}
  - Record if the two tosses matched {Yes, No}

Definition 35. An event is a set of outcomes of a random trial, or in other words, a subset of the sample space.

There are different types of events.

- An *impossible event* refers to an outcome which is actually not possible, that is, which does not belong to the sample space. For example, if you roll a dice, the number that you will get belongs to the sample space $\{1, 2, 3, 4, 5, 6\}$, therefore, roll a dice and obtain 7 is an impossible event.
- A *certain event* is on the contrary an event which always happens, which corresponds to the whole sample space, such as: roll a dice, and get a number which is less than 10. Any number in the sample space $\{1, 2, 3, 4, 5, 6\}$ is less than 10.
- Two events are said to be *mutuall exclusive* when they cannot happen at the same time. For example, you roll a dice, the events "get an even number" and "get a number divisible by 5" are mutually exclusive, they cannot happen both at the same time.

We now want to define the notion of *probability of an event*. Informally, this represents the likelihood that an event will occur, or the ratio of the number of wanted outcomes, by the number of possible outcomes. For example, if you toss a fair coin, the probability of a head (that is, the likelihood that a head happens), is 1/2. This comes from the fact that you have two possible outcomes, head and tail, and both of them are equally likely (so 2 at the denominator). On the numerator, you want a head (so 1 at the numerator).

You may also want to repeat a given random experiment, say n times. When you look at a particular event E, over the n times, it will appear n_E times. Therefore the *frequency* of occurence of E over these n trials is

$$f_E = \frac{n_E}{n}.$$

The notion of frequency is different from that of probability of an event, but it is also related. For example, if you toss a coin 10 times, and you are interested in counting the number of occurences of the event $E = \text{head}$, maybe you get $n_E = 6$, therefore $f_E = 6/10$.

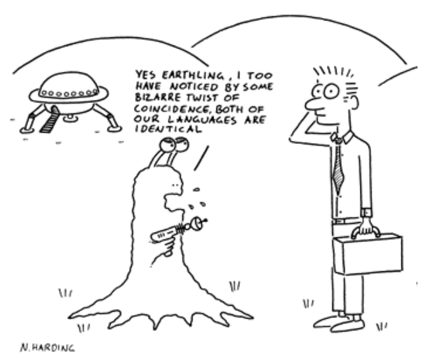
Events

An event is a set of outcomes of a random trial (a subset of the sample space).

- Impossible event : event not in the sample space.
  - *E.g Roll a dice, get more than 6 on top.*
- Certain event: event which is the whole sample space.
  - E.g. *Roll a dice, get a number less than 10.*
- Mutually exclusive events: events which cannot happen at the same time.
  - *E.g. Roll a dice, get an even number which is divisible by 5.*

Probability Of Events

- The likelihood that an event will occur
  - When tossing a fair coin, the probability of a head is 0.5.
- Empirical interpretation
  - Repeat an experiment is n times
  - An event E occurs n_E times
  - Then $f_E = n_E / n$ is the frequency of occurrence of E in n trials
    - such a frequency measure is not necessarily a proof of the probability of event E, but can be an "indicator"



The frequency is then slightly different that what you may have expected, since with a probability of 1/2 and 10 trials, you may have expected the number of heads ot be 5. Therefore the frequency is an "indicator".

We next introduce a formal definition of probability.

Definition 36. A probability space is a sample space A, with some events $E_i \in A$, and probability measure P, satisfying 3 axioms:

1. $0 \leq P(E_i) \leq 1$ for every event $E_i \in A$.
2. $P(A) = 1$ and $P(\varnothing) = 0$.
3. $P(E_1 \cup E_2 \cup \ldots \cup E_k) = P(E_1) + P(E_2) + \ldots + P(E_k)$ when $E_i \cap E_j = \varnothing$, $i \neq j$.

Example 55. Suppose that the random trial is tossing a fair coin: The sample space is { head, tail }. Because the coin is fair (unbiased), head and tail are equally likely events (P(head)=P(tail)), and they are mutually exclusive (P(head)+P(tail)=1). Therefore, we obtain formally what our intuition told us, namely, that P(head)=P(tail)=1/2.

More generally, if there are n equally likely mutually exclusive (and spanning the sample space) events $E_1, \ldots, E_n$, then we get: $P(E_1) = \cdots = P(E_n)$ and $P(E_1) + \ldots + P(E_n) = 1$ therefore $P(E_i) = 1/n$ for all every event E_i.

Example 56. Suppose that you choose 4 cards from a deck of 52 cards. What is the probability of getting 4 kings? There are $C(52, 4)$ ways of choosing 4 cards. Now all are equally likely, but only one of these choices has all four kings. Therefore the probaiblity of getting all 4 kings is

$$\frac{1}{C(52, 4)}.$$

Axioms For A Probability Space

For a sample space A, *and some* event $E_i \in A$:

- Axiom 1: $0 \leq p(E_i) \leq 1$ for every event $E_i \in A$
- Axiom 2: $P(A) = 1$ and $P(\phi) = 0$

  (Event A must happen everytime the experiment is done since every event belongs to A)
- Axiom 3: $P(E_1 \bigcup E_2 \bigcup ... \bigcup E_k) = P(E_1) + P(E_2) + ... + P(E_k)$

  when $E_i \bigcap E_j = \phi$ for $i \neq j$

  (i.e., E_i and E_j are *mutually exclusive)*

Equally Likely Outcomes: Symmetry

- Example: Tossing an unbiased coin
  - Sample space is {H,T} (H=head, T=tail)
  - No reason for H T to occur more often than T: P(H)=P(T)
  - H and T are mutually exclusive events: P(H)+P(T)=1
  - Thus P(H)=P(T)=0.5
- In general: if there are *n* equally likely mutually exclusive (and spanning the sample space) events, then the elementary probability of each such event is **1/n**

Exercises for Chapter 5

Exercise 46. A set menu proposes 2 choices of starters, 3 choices of main dishes, and 2 choices of desserts. How many possible set menus are available?

Exercise 47. Consider the set $A = \{1, 2, 3\}$, $P(A)$ =power set of A.

- Compute the cardinality of $P(A)$ using the binomial theorem approach.
- Compute the cardinality of $P(A)$ using the counting approach.

Exercise 48.
- If you toss two coins, what is the probability of getting 2 heads?
- If you toss three coins, what is the probability of getting exactly 2 heads?

Exercise 49. Ten fair coins are tossed together. What is the probability that there were at least seven heads?

Exercise 50. Snow white is going to a party with the seven dwarves. Each of the eight of them owns a red dress and a a blue dress. If each of them is likely to choose either colored dress randomly and independently of the other's choices, what is the chance that all of them go to the pary wearing the same colored dress?

Examples for Chapter 5

Let us recall one storage application, given in Example 8.

Example 57. Suppose you want to store 200GB of data, and the shop is selling disks of 100 GB each. Then you can buy 4 disks, store half of your data (let us call it D_1) on disk 1, the other half (say D_2) on disk 2, then copy the content of disk 1 to disk 3, and the content of disk 2 to disk 4. We get thus the following data allocation:

$$\text{disk } 1 : D_1,\ \text{disk } 2 : D_2,\ \text{disk } 3 : D_1,\ \text{disk } 4 : D_2.$$

This strategy does tolerate any one disk failure. It does not tolerate any two disks failures, as we already know. However, what is the probability of actually losing the data in case of two disks failures? The number of patterns with 2 failures is $C(4, 2) = 6$, while the number of patterns creating a data loss is 2, therefore the probability is

$$\frac{2}{6} = \frac{1}{3}.$$

Our next example is a famous puzzle, **the Hat Problem.** It is famous because it made the news!

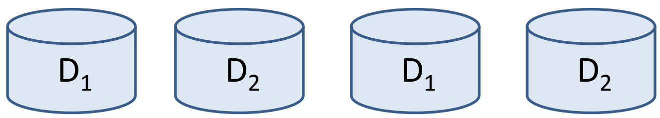
`http://www.nytimes.com/2001/04/10/science/`
`why-mathematicians-now-care-about-their-hat-color.html`

Here is how the puzzle goes: N players enter a room. A red or blue hat is placed on each person's head. The color of each hat is determined by a fair coin toss, that is P(blue)=P(red), with the outcome of one coin toss having no effect on the others. Each person can see the other players' hats but not his own.

- The players must simultaneously guess the color of their own hats or pass.
- The group shares a prize if at least one player guesses correctly, and no player guesses incorrectly.
- No communication of any sort is allowed, exepct for any initial strategy session before the game begins. For a strategy to be acceptable, it must always result in at least one prisoner making a guess.

Example (II)

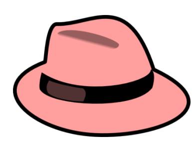
$D = (D_1, D_2)$



- If one hard disk fails, your data is safe.
- What is the probability of losing your data in case two hard disks fail?

$$\frac{2}{C(4,2)} = \frac{2}{6} = \frac{1}{3}$$

The Hat Problem



N players enter a room

- A red or blue hat is placed on each person's head.
  - P(red)=P(blue)=1/2, independently.
- Each player sees the other hats but not his own.
  - The players must simultaneously guess the color of their own hats or pass.
- Win *if at least one player guesses correctly and no players guess incorrectly*.
  - No communication is allowed, except for any initial strategy session before the game begins.

Example 58. If $N = 1$, there is only one player, thus all he can do is make a guess, the probability of winning is simply $1/2$.

Example 59. If $N = 2$, there are two players. If both players guess randomly, their chance of winning is only $1/4$, because they try to guess one combination of colors among 4 possible combinations. A single player only making a guess is better, the probability of winning is then $1/2$.

Suppose now that $N = 3$. Let us assign the number 0 to a hat of blue color, and 1 to a hat of red color. There are 8 possible hat assignements:

$$000, 100, 010, 110, 001, 101, 011, 111.$$

Therefore, if a player sees two hats of the same color, he guesses the opposite color (this is an acceptable strategy, there must always be two hats of the same color). Otherwise he passes. This corresponds to

$$100, 010, 110, 001, 101, 011,$$

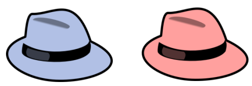
where only one player will see two hats of the same color, will guess his color to be the other one, and will be right with probability $6/8 = 3/4$. The game will be won with probability $3/4 > 1/2$.

In fact, we can prove that this strategy is optimal!

(1) First, we need to note that the number of correct guesses and the number of incorrect guesses are the same. The reason is because the probability of making a correct/incorrect guess is $1/2$. For example, when $N = 2$, when both players are making a guess, there is one chance of winning (2 correct guesses), and three chances of losing (2 wrong guesses, 1 wrong guess and 1 correct guess, twice). Thus the total number of correct guesses is 4, and the number of incorrect guesses is 4 as well! For another example, when $N = 3$, the proposed strategy comprises 6 correct guesses (one for each of the 6 wins), and 6 wrong guesses (3 incorrect guesses per each loss, and there are two losses). More precisely, for the correct guesses 100, 010, 001, correspond the 3 incorrect guesses 000, and for the correct guesses 011,101,110, correspond the 3 incorrect guesses 111.

(2) Suppose we could get a better strategy, where we get 7 wins instead of 6, and thus 1 loss. For a winning strategy, no incorrect guess can be made, therefore we need at least 7 right guesses, which means in turn, using the above argument, 7 wrong guesses. But this is not possible to have 7 wrong guesses in one loss, and only 3 players.

The Hat Problem



0 1

000 100 010 110 001 101 011 111

Strategy: If the other two guys have the same hat color, "guess the opposite", if they have different colors, stay silent!

- *Chance of winning with this strategy*: 3/8+3/8=0.75

The Hat Problem

- Optimal strategy?
- Number of correct guesses = number of incorrect guesses
- Better strategy: 7 wins & 1 loss
- At least 7 correct guesses, impossible to have 7 incorrect guesses in one loss and 3 players