

# Chapter 1

## Probabilities

A lot of the times in our daily life, we can not predict the outcome of an event exactly, because of the complicated situations and the knowledge we don't know. However if this event is repeated many times, we may be able to calculate the "likelihood" of certain outcome, and this "likelihood" is called *probability*.

## 1.1 Probability and simulation

### Definition of probability

When throw a coin, we say the probability of getting a *head* is 0.5. What does this mean? One way to interpret this is that: if we throw a fair coin many many time, half of the time will come up with a head. This is the **frequentists'** view of probability. The other way to interpret is: If a coin is tossed, the strength of your belief that it will come up with a *head* is 0.5 if the strength of your belief is measured by numbers between 0 and 1. This is the **Baysian** view of probability.

In this book and the AP exam we hold to the frequentists' view of probability.

The **probability** of an outcome of a trial is defined as the relative frequency of this outcome if the trial is repeated many many times.

If we want know the probability of getting a head when tossing a fair coin, we toss the coin millions of billions of times and calculate the ratio of the number of heads to the total number of trials.

Now the problem is how do we know that the *relative* frequency converges as the number of trials increases. The **law of large numbers** guarantees the convergence, which says the limit of the relative frequency exists as the number of trials increases.

### Simulation

The definition of the probability gives us a way to calculate the probability simply by repeating the trials. If you want find out the probability of getting a head, just tossing a coin many many times and calculate the relative frequency. However this is time consuming and sometimes not practical. We can simulate the process of tossing coins.

We can use a random number generator or table of random digits to generate random numbers between 0 and 9, if the number generated is between from 0 to 4, we take it as head, otherwise as tail. We generate the random numbers many many times and calculate the relative frequency of numbers from 0 to 4. This is more like an **experimentation** of probability. We call this process **simulation**.

### **Golden Ticket Parking Lottery**

At a local high school, 95 students have permission to park on campus. Each month, the student council holds a golden ticket parking lottery at a school assembly. The two lucky winners are given reserved parking spots next to the schools main entrance. Last month, the winning tickets were drawn by a student council member from the AP<sup>®</sup> Statistics class. When both golden tickets went to members of that same class, some people thought the lottery had been rigged. There are 28 students in the AP Statistics class, all of whom are eligible to park on campus. Design and carry out a simulation to decide whether its plausible that the lottery was carried out fairly.

## 1.2 Probability models

### A little bit of counting

#### 1. Principle of multiplication

Ways from **A** to **C** is shown in figure 1.1, there are three ways from **A** to **B**, and there are two ways from **B** to **C**. And, no matter which way is chosen from **A** to **B**, there are always 2 ways to choose from **B** to **C**.

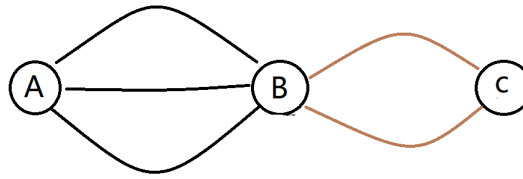


Figure 1.1: Principle of multiplication

To find the total number of possible ways from **A** to **C**, just do multiplication.

$$\text{Total number of ways} = 3 \times 2 = 6$$

The above is called *principle of multiplication*.

How many subsets are there for a set with  $n$  elements?

#### 2. Permutation

**Permutation** is to arrange a number of distinct elements in *order*. For example, we want selected 3 out of 6 and arrange them in order. How many possible arrangements are there?

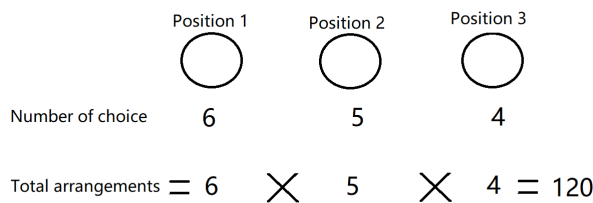


Figure 1.2: Permutation

As shown in figure 1.2, first we fix three positions, then we fill out those positions one by one. There are 6 choices for position 1, 5 for position 2 and 4 for position 3. And no matter what choice is made in position 1, the number of choices for position 2 is the same. The same situation is true for position 2 and position 3. Thus, by applying the *multiplication principle* the total number of arrangements is given by

$$6 \times 5 \times 4 = 120.$$

The above number of arrangement can be represented by notation  ${}_3\mathbf{P}_6$ , where  $\mathbf{P}$  means *permutation*.

$${}_m\mathbf{P}_n = n \times (n - 1) \times \cdots \times (n - m + 1)$$

The total arrangements of arranging  $n$  distinct elements in order can be represented by notation  $n!$ , read as **n factorial**.

$$n! = n \times (n - 1) \times \cdots \times 2 \times 1$$

### 3. Combination

Select  $m$  elements out of  $n$  distinct elements, without arranging them in order, is called **combination**. The notation is  ${}_m\mathbf{C}_n$ .

Select 3 out of 6 elements and arrange the selected 3 in order. This is a permutation problem and the solution is given by  ${}_3\mathbf{P}_6$ . Now lets solve this problem in another way. We divide the process into two steps. First, select 3 out of 6, which is a *combination* problem, and the total combinations is denoted as  ${}_3\mathbf{C}_6$ . Second, arrange the selected 3 in order. No matter how the 3 elements in selected in the first step, there will always be  $3!$  arrangements for the second step. Apply the *multiplication principle*.

$${}_3\mathbf{P}_6 = {}_3\mathbf{C}_6 \times 3!$$

. Thus,

$${}_3\mathbf{C}_6 = \frac{{}_3\mathbf{P}_6}{3!} = \frac{6 \times 5 \times 4}{3!} = \frac{6!}{3!3!}$$

Similarly

$${}_m\mathbf{C}_n = \frac{n!}{m!(n - m)!}$$

**Prove:**  ${}_m\mathbf{C}_n = {}_{n-m}\mathbf{C}_n$

## Probability models

Toss a coin, if it comes up with a head, we denote it as "H". If it is tail, it is denoted as "T". We toss a coin twice, and order matters. All possible outcomes can be given by the following set.

$$\mathbf{S} = \{(H,H), (H,T), (T,H), (T,T)\}$$

The set of all possible outcomes is called the **sample space**, often denoted as **S**. A **probability model** consists of the sample space and the probabilities assigned to each outcome.

An **event** is a subset of the sample space. For example event **A** can be given by

$$\mathbf{A} = \{(H, H), (H, T), (T, H)\}$$

Clearly  $\mathbf{A} \subset \mathbf{S}$ , and event **A** can be interpreted as "tossing a coin twice, and get at least one head". Operations of sets can be applied to events.  $\mathbf{P}(\mathbf{A} \cup \mathbf{B})$  means the probability that either **A** or **B** happens.  $\mathbf{P}(\mathbf{A} \cap \mathbf{B})$  means the probability that both **A** and **B** happen.

If all possible outcomes are equally likely, then we say this probability model is a **classical probability model**.

When tossing a coin twice, and  $\mathbf{S} = \{(H,H), (H,T), (T,H), (T,T)\}$ , each outcome has a probability 0.25. It is a *classical probability model*.

If we toss a coin twice, and the sample space is given by  $\mathbf{S} = \{(H,H), (H,T), (T,T)\}$ . In this case we assign probability 0.5 to outcome  $(H, T)$ , which means we don't distinguish the order of the tossing,  $(H, T)$  and  $(T, H)$  are the same outcome. In this case, it is not a *classical probability model*.

For classical probability model, the probability of an event **A** can be given by

$$\mathbf{P}(\mathbf{A}) = \frac{n(\mathbf{A})}{n(\mathbf{S})},$$

where  $n(\mathbf{A})$  means the number of outcomes in **A**, and  $n(\mathbf{S})$  means the total number of outcomes.

There will either be an earthquake or not, only two outcomes. Thus the probability of having an earthquake tomorrow is 1/2. What is wrong with this argue?

### Exercise:

Toss a die twice. Find the following probabilities.

- Event **A**: the sum is divisible by 2.
- Event **B**: the sum is divisible by 3.
- Event **C**: the sum is divisible by 5.
- $\mathbf{A} \cup \mathbf{B}$ . Interpret the meaning of this event.
- $\mathbf{A} \cap \mathbf{B}$ . Interpret the meaning of this event.
- $\mathbf{A} \cap \mathbf{C}$ . Interpret the meaning of this event.

## 1.3 Probability rules

### General addition rule

Since events are subsets of sample space  $\mathbf{S}$ , there relations can be represented by the *Venn Diagram* as shown in figure 1.3. Lets simply take the probabilities of different events as areas enclosed by different events, and  $\mathbf{P}(\mathbf{S}) = 1$ .

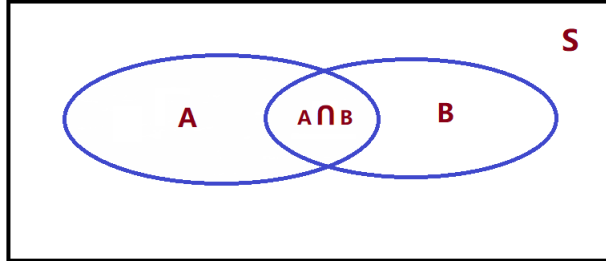


Figure 1.3: General addition rule

According to the *Venn Diagram*, we have

$$\mathbf{P}(\mathbf{A} \cup \mathbf{B}) = \mathbf{P}(\mathbf{A}) + \mathbf{P}(\mathbf{B}) - \mathbf{P}(\mathbf{A} \cap \mathbf{B}).$$

The above is called the **general addition rule**. If  $\mathbf{P}(\mathbf{A} \cap \mathbf{B}) = 0$ , we say  $\mathbf{A}$  and  $\mathbf{B}$  are **mutually exclusive**, which means they can not happen at the same time.

If  $\mathbf{A}$  and  $\mathbf{B}$  are *mutually exclusive*, then

$$\mathbf{P}(\mathbf{A} \cup \mathbf{B}) = \mathbf{P}(\mathbf{A}) + \mathbf{P}(\mathbf{B}).$$

$\mathbf{A}^C$  is the **complementary** of  $\mathbf{A}$ ,  $\mathbf{P}(\mathbf{A} \cap \mathbf{A}^C) = \mathbf{P}(\emptyset) = 0$ . Thus

$$\mathbf{P}(\mathbf{S}) = \mathbf{P}(\mathbf{A}) + \mathbf{P}(\mathbf{A}^C) - \mathbf{P}(\mathbf{A} \cap \mathbf{A}^C) = 1.$$

#### Exercise:

Toss a die twice.

Event  $\mathbf{A}$  means "the sum is divisible by 2".

Event  $\mathbf{B}$  means "the sum is divisible by 3".

Event  $\mathbf{C}$  means "the sum is divisible by 5".

(a) Which of the following pairs of events are mutually exclusive?

(1)  $\mathbf{A}$ ,  $\mathbf{B}$       (2)  $\mathbf{A}$ ,  $\mathbf{C}$       (3)  $\mathbf{B}$ ,  $\mathbf{C}$ .

(b) Find  $\mathbf{P}(\mathbf{A} \cup \mathbf{B})$ ,  $\mathbf{P}(\mathbf{A} \cup \mathbf{C})$ ,  $\mathbf{P}(\mathbf{B} \cup \mathbf{C})$ .

## General multiplication rule

Take a look at this example. Randomly select 178 students, and check whether they have got their ear pierced. The results is give by figure 1.4.

Pierced Ears	Gender		Total
	Male	Female	
Yes	19	84	<b>103</b>
No	71	4	<b>75</b>
<b>Total</b>	<b>90</b>	<b>88</b>	<b>178</b>

Figure 1.4: Data of pierced ears

Let's denote event **A** as "male" and event **B** as "pierced ears". For a randomly chosen students, the probability that his/her ears are piecered is

$$P(B) = \frac{103}{178}.$$

If we tweak the situation a little bit and ask "For a randomly chosen student, we know he is male, what is the probability his ears are pierced?" Since we already know this student is male, we only have to consider the probability among the group of male students. In this male group, the probability that his ears are pierced is  $\frac{19}{90}$ . We call this kind of probability **conditional probability**, and denoted as  $P(B|A)$ , which means the probability of **B**("pierced ears") on condition that **B** happened("male"). We have the following formulae:

$$P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B).$$

calculate  $P(A|B)$   
and interpret its mean-  
ing.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

Those formulae are very important. They play a key role in **Baysian statistics** as well.

If  $P(B|A) = P(B)$ , we say **A** and **B** are **independent**, which means whether **A** happens or not, it won't influence the probability of **B**. In fact

$$P(B|A) = P(B) \implies P(A|B) = P(A)$$

This means *independence* is *mutual*. If **A** is independent of **B**, then **B** is independent of **A**.

**Prove:**  $P(B|A) = P(B) \implies P(A|B) = P(A)$



If **A** and **B** are independent, then

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \cdot P(\mathbf{B})$$

. This formula can be generalized to more than two independent events.

### **The Challenger Disaster?**

On January 28, 1986, Space Shuttle Challenger exploded on takeoff. All seven crew members were killed. Following the disaster, scientists and statisticians helped analyze what went wrong. They determined that the failure of O-ring joints in the shuttles booster rockets was to blame. Under the cold conditions that day, experts estimated that the probability that an individual O-ring joint would function properly was 0.977. But there were six of these O-ring joints, and all six had to function properly for the shuttle to launch safely.

Assuming that O-ring joints succeed or fail independently, find the probability that the shuttle would not launch safely under similar conditions.

**Tree diagram and conditional probability****Who Visits *YouTube*?**

Video-sharing sites, led by YouTube, are popular destinations on the Internet. Lets look only at adult Internet users, aged 18 and over. About 27% of adult Internet users are 18 to 29 years old, another 45% are 30 to 49 years old, and the remaining 28% are 50 and over. The Pew Internet and American Life Project finds that 70% of Internet users aged 18 to 29 have visited a video-sharing site, along with 51% of those aged 30 to 49 and 26% of those 50 or older.

Suppose we select an adult Internet user at random.

1. Find the probability that this person has visited a video-sharing site. Show your work.
2. Given that this person has visited a video-sharing site, find the probability that he or she is aged 18 to 29. Show your work.