

# ENGG2760 Probability for Engineers

Ryan Chan

December 13, 2024

### **Abstract**

This is a note for **ENGG2760 - Probability for Engineers** for self-revision purpose ONLY. Some contents are taken from lecture notes and reference book.  
Mistakes might be found. So please feel free to point out any mistakes.

# Contents

<b>1</b>	<b>Probability and Counting</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.2	Permutation and Combination . . . . .	3
<b>2</b>	<b>Probability Models and Axioms</b>	<b>6</b>
2.1	Basic Definitions . . . . .	6
2.2	Probability Axioms . . . . .	7
2.3	Rules for Probability Calculation . . . . .	7
<b>3</b>	<b>Conditional Probability and Bayes' Rule</b>	<b>9</b>
3.1	Conditional Probability . . . . .	9
3.2	Bayes' Rule . . . . .	12
3.3	Independence . . . . .	13
<b>4</b>	<b>Random Variables</b>	<b>16</b>
4.1	Introduction . . . . .	16
4.2	Binomial Random Variable . . . . .	17
4.3	Geometric Random Variable . . . . .	18
4.4	Cumulative Distribution Function (CDF) . . . . .	20
4.5	Poisson Random Variable . . . . .	21
4.6	Properties of Random Variables . . . . .	22
<b>5</b>		<b>25</b>
5.1	. . . . .	25
<b>6</b>		<b>26</b>
6.1	. . . . .	26
<b>7</b>		<b>27</b>
7.1	. . . . .	27
<b>8</b>		<b>28</b>
8.1	. . . . .	28
<b>9</b>		<b>29</b>
9.1	. . . . .	29
<b>10</b>		<b>30</b>
10.1	. . . . .	30

# Chapter 1

## Probability and Counting

### 1.1 Introduction

We will start with some basic definitions.

**Definition 1.1.1 (Sample Space).** The sample space  $\Omega$  is the set of all possible outcomes

For example, when flipping three coins, we have  $2^3 = 8$  outcomes:

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

**Definition 1.1.2 (Event).** An event is a subset of the sample space.

Following the above example, if  $A$  is the event that at least two heads occur, we have:

$$A = \{HHH, HHT, HTH, THH\}$$

**Definition 1.1.3.** The probability of an event is the sum of the probability of its outcomes.

- Probabilities are non-negative.
- Probabilities add up to one.

Again from the above example, we see that the probability of each event is equal to  $\frac{1}{8}$ , and they can be summed up to 1.

**Definition 1.1.4.** The probability of an event is the sum of the probabilities of its outcomes.

For event  $A$ , the probability would be

$$\mathbb{P}(A) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}$$

**Proposition 1.1.1 (Uniform Probability Law).** If the outcomes in  $\Omega$  are equally likely, then the probability of event  $A$  will be

$$\mathbb{P}(A) = \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } \Omega} = \frac{|A|}{|\Omega|}$$

**Remark.** It can only be used when every outcome is equally likely.

---

For event  $A$ , the probability would be

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}$$

**Example.** We roll two dice. Which of the following outcome is more likely for the sum of the two dice?

1. 11
2. 12
3. equally likely

**Solution:** For the sum to be 11, we can have (5, 6) and (6, 5). However, for the sum to be 12, we can only have (6, 6). Therefore, for  $|\Omega| = 6^2 = 36$ ,

$$\mathbb{P}(11) = \frac{2}{36}, \quad \mathbb{P}(12) = \frac{1}{36}$$

Therefore, the sum of 11 would be more likely to occur.

## 1.2 Permutation and Combination

### 1.2.1 Counting via Product Rule

**Proposition 1.2.1 (Product Rule).** Suppose there are  $n$  possible outcomes for Experiment 1 and  $m$  possible outcomes for Experiment 2, where the two experiments are independent. Then, there are  $m \times n$  possible outcomes for the two experiments.

For example, when flipping three coins, each of them has two possible outcomes. Therefore, there are in total  $2 \times 2 \times 2 = 8$  possible outcomes.

We can then generalize this rule for cases that the outcomes of experiment 1 may affect the outcomes of experiment 2.

**Proposition 1.2.2 (Generalized Product Rule).** Suppose that

- There are  $n$  possible outcomes for Experiment 1.
- For every outcome of Experiment 1, there are  $m$  possible outcomes for Experiment 2.

Then, there are  $m \times n$  possible outcomes for the two experiments.

For example, when finding all possible outcomes for rolling two dice with different values, the outcomes of the first experiment, i.e. rolling the first die, would be 6. The outcomes of the second experiment, i.e. rolling the second die, would be 5 (since we need to exclude the outcome of the first die). Then, there are in total  $6 \times 5 = 30$  possible outcomes.

**Example.** We roll two dice. What is the probability that they come out with different values?

**Solution:** Let  $A$  be the desired event. Then we have

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{6 \times 5}{6 \times 6} = \frac{5}{6}$$

**Example.** We roll two dice. What is the probability that the sum of dice equals 7? What is the probability that the sum of dice is an odd number?

**Solution:** Let  $A$  be the event that the sum of dice equals 7. Then we have

$$A = \{(1, 6), (2, 5), \dots, (6, 1)\}$$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{6^2} = \frac{1}{6}$$

Let  $B$  be the event that the sum of dice is an odd number. Then we have

$$B = \{(1, 2), (1, 4), \dots, (6, 5)\}, \quad |B| = 6 \times 3,$$

where for each number in the first die, there will be exactly three numbers in the second die that can be added up to an odd number. Thus,

$$\mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{6 \times 3}{6^2} = \frac{1}{2}$$

**Example.** We again roll two dice. What is the probability that the first die is bigger than the second die?

**Solution:** In this case, we cannot use generalized product rule since for every outcome in the first experiment, there will be a different outcome in the second experiment. Let  $A$  be the desired event. Then we have

$$A = \{(2, 1), (3, 1), \dots, (6, 5)\}$$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{15}{6^2} = \frac{5}{12}$$

## 1.2.2 Permutation

**Definition 1.2.1 (Permutation).** A permutation of  $n$  different objects is an arrangement of the objects into an ordered sequence (order matters).

**Proposition 1.2.3.** For  $n$  different objects, there exists  $n!$  different permutations:

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

**Example.** We roll six dice. How many ways are there for the six dice to have different values? What is the probability of that event?

**Solution:** Let  $A$  be the desired event. Then we have

$$|A| = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 6! = 720, \quad \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{6!}{6^6}$$

**Example (Birthday Paradox).** Suppose there are  $n$  people in a room. We assume that a year only has 365 days, and that every day is equally likely to be the birthday of a person. What is the probability that at least two people have the same birthday? Here we assume that  $n < 365$ .

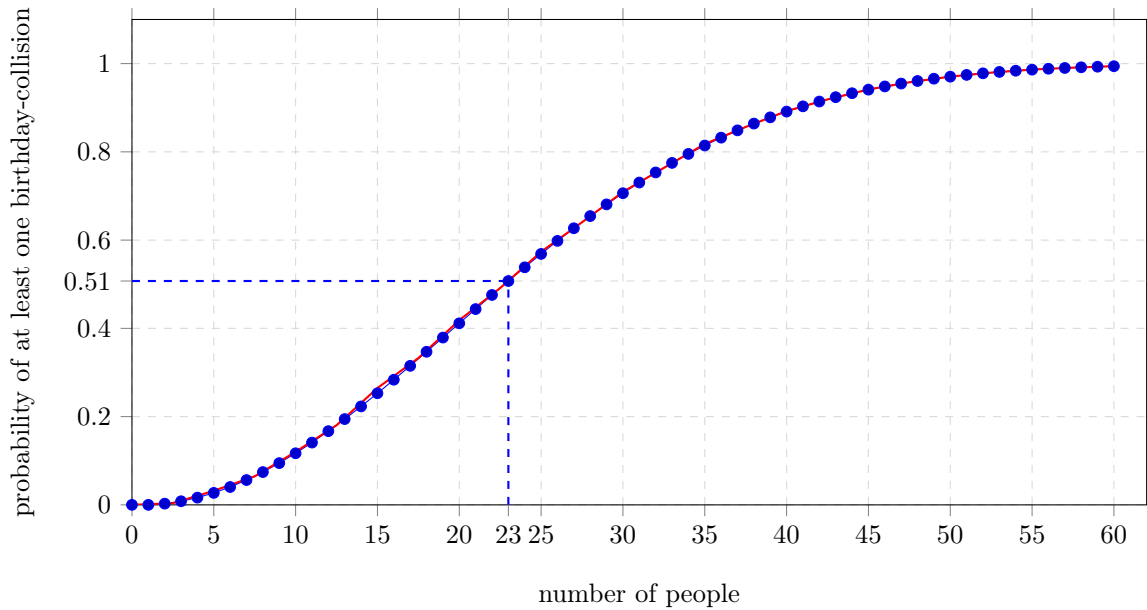
For sample space  $S$  we have the set of all possible sequences of  $n$  birthday, the  $|S| = 365^n$ .

Let  $T$  be the event in which at least two birthdays are the same. Then we have

$$\mathbb{P}(T) = 1 - \frac{365 \times 364 \times \dots \times (365 - n + 1)}{365^n},$$

where the term  $(365 \times 364 \times \dots \times (365 - n + 1))$  is to count the possible outcomes for the event that all birthdays are distinct.

Birthday paradox could be visualized as below:



Adapted from [MartinThoma](#)

### 1.2.3 Binomial Coefficient

**Proposition 1.2.4** (Binomial Coefficient or " $n$ -Choose- $k$ "). Given a set  $S$  of size  $n$ , the number of subsets of size  $k$  will be

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

It can also be understood as the number of possible arrangements of  $k$  objects of Type A and  $n-k$  objects of Type B into an ordered sequence.

**Example.** A box contains 8 red balls and 2 blue balls. You draw 2 balls at random (without replacement). What is the probability that the two balls have different colors?

**Solution:** Let  $A$  be the desired event.

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{\binom{8}{1}\binom{2}{1}}{\binom{10}{2}} = \frac{16}{45}$$

**Proposition 1.2.5** (Multinomial Coefficient). For a set  $S$  of size  $n$ , the number of partitioning of the set to partitions of size  $k_1, k_2, \dots, k_t$  (noted that  $n = k_1 + k_2 + \dots + k_t$ ) will be

$$\binom{n}{k_1, k_2, \dots, k_t} = \frac{n!}{k_1! k_2! \dots k_t!}$$

It can also be understood as the number of possible permutations of  $k_1$  objects of Type 1,  $k_2$  objects of Type 2, ..., and  $k_t$  objects of Type  $t$ .

## Chapter 2

# Probability Models and Axioms

### 2.1 Basic Definitions

We will introduce some definitions here as well.

**Definition 2.1.1 (Complement).** The complement of event  $A$  (denoted by  $A^c$ ) is the opposite event of  $A$ . In other words,  $A^c$  happens if and only if  $A$  does not happen.

Again, when flipping three coins, we have the following sample space:

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

Let  $A$  be the event that at least two heads occur. Then for  $A^c$ , we have:

$$A^c = \{TTT, HTT, THT, TTH\}$$

**Definition 2.1.2 (Intersection of Events).** The intersection of events happens when all the events occur. We denote this intersection of event  $A$  and  $B$  with  $A \cap B$ .

Let  $B$  be the event that no consecutive heads occurs. Then, for  $A \cap B$ , we have the event that at least two heads and no consecutive heads occur.

$$A \cap B = \{HTH\}$$

**Definition 2.1.3 (Union of Events).** The union of events happens when at least one of the events occur. We denote the union of events  $A$  and  $B$  with  $A \cup B$ .

For example, for  $A \cup B$  in the above example, we have

$$A \cup B = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

**Definition 2.1.4 (Disjoint Events).** We call event  $A_1, A_2, \dots$  disjoint events (or mutually exclusive events) if the intersection of every two events  $A_i, A_j (i \neq j)$  is the null event:

$$\forall i \neq j : A_i \cap A_j = \emptyset$$

Let  $C$  be the event that at least three heads occur. Then

$$B \cap C = \emptyset.$$



## 2.2 Probability Axioms

**Definition 2.2.1 (Axioms of Probability).** A probability assignment  $\mathcal{P}$  to sample space  $\Omega$  should satisfy the following three axioms:

1. For every event  $A$ ,  $0 \leq \mathbb{P}(A)$ ;
2.  $\mathbb{P}(\Omega) = 1$ ;
3. If event  $A_1, A_2, \dots$  are disjoint,  $\mathbb{P}(A_1 \cup A_2 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots$

Follow these axioms, and we can prove most of the rules for probability calculation.

## 2.3 Rules for Probability Calculation

**Proposition 2.3.1 (Complement Rule).** For every event  $E$  and its complement  $E^c$ :

$$\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$$

**Proposition 2.3.2 (Difference Rule).** If event  $E, F$  satisfy  $E \subseteq F$ , then:

$$\mathbb{P}(F \cap E^c) = \mathbb{P}(F) - \mathbb{P}(E)$$

**Remark.** As a result, if  $E \subseteq F$ , then  $\mathbb{P}(E) \leq \mathbb{P}(F)$

**Proof.**

$$\mathbb{P}(F \cap E^c) = \mathbb{P}(F) - \mathbb{P}(E)$$

$$\mathbb{P}(F) = \mathbb{P}(F \cap E^c) + \mathbb{P}(E)$$

Since  $(F \cap E^c) \cap E = F \cap (E^c \cap E) = F \cap \emptyset = \emptyset$ ,  $\mathbb{P}(F \cap E^c) + \mathbb{P}(E) \Rightarrow (F \cap E^c) \cup E$

$$(F \cap E^c) \cup E = (F \cup E) \cap (E^c \cup E)$$

$$(F \cap E^c) \cup E = (F \cup E) \cap \Omega$$

$$(F \cap E^c) \cup E = F \cup E$$

$$(F \cap E^c) \cup E = F \quad (\text{for } E \subseteq F)$$

■

**Proposition 2.3.3 (Inclusion-Exclusion Principle).** For events  $E, F$ :

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$$

**Remark.** We can generalize the principle to more than two events. For example,

$$\mathbb{P}(E \cup F \cup G) = \mathbb{P}(E) + \mathbb{P}(F) + \mathbb{P}(G) - \mathbb{P}(E \cap F) - \mathbb{P}(E \cap G) - \mathbb{P}(F \cap G) + \mathbb{P}(E \cap F \cap G)$$

**Example.** In a city, 10% of the people are rich, 5% are famous, and 3% are both rich and famous. For a randomly-selected person in the city, find the probability for the following Events.

Here we let  $R$  be the event that the person is rich,  $F$  be the event that the person is famous,

1. The person is not rich.

$$\mathbb{P}(R^c) = 1 - \mathbb{P}(R) = 1 - 0.1 = 0.9$$

- 
2. The person is not rich but is famous.

$$\mathbb{P}(R^c \cap F) = \mathbb{P}(F) - \mathbb{P}(F \cap R) = 0.05 - 0.03 = 0.02$$

3. The person is neither rich nor famous.

$$\mathbb{P}(F^c \cap R^c) = 1 - \mathbb{P}(F \cup R) = 1 - \mathbb{P}(F) - \mathbb{P}(R) + \mathbb{P}(F \cap R) = 1 - 0.05 - 0.1 + 0.03 = 0.88$$

## Chapter 3

# Conditional Probability and Bayes' Rule

### 3.1 Conditional Probability

Let's begin with an example

**Example.** We toss 3 coins. You win if at least two heads come out. What is the probability of winning?

**Solution:**

$$\Omega = \{HHH, HHT, \dots, TTT\} \Rightarrow \mathbb{P}(W) = \frac{|W|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}$$

However, if it is given that the first coin was tossed and came out head. How does this affect your chances of winning? Since the sample space now changes due to the condition that the first toss is heads, the way we find the probability differs.

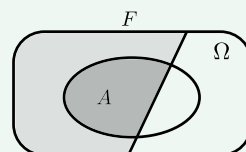
$$\Omega' = \{HHH, HHT, HTH, HTT\}, \quad W' = \{HHH, HHT, HTH\}$$

This give the probability  $\frac{3}{4}$ .

#### 3.1.1 Conditional Probability

**Definition 3.1.1** (Conditional Probability).

The Conditional Probability  $\mathbb{P}(A|F)$  represents the probability of event  $A$  assuming (or given) that event  $F$  happened.



**Remark.** All the outcomes of  $\Omega$  and event  $A$  in  $F$  should be excluded in the calculation.

The conditional probability of  $A$  with respect to reduced sample space  $F$  is given by the formula:

$$\mathbb{P}(A|F) = \frac{\mathbb{P}(A \cap F)}{\mathbb{P}(F)}$$

**Example.** You roll two dice. You win if the sum of the outcomes is 8. If the first die toss is a 4, should you be happy?

**Solution:** Considering the initial case, we have:

$$\mathbb{P}(W) = \frac{|W|}{|\Omega|} = \frac{5}{36}$$

Let  $D$  be the event that the first toss is a 4. Given  $D$ , we have:

$$\mathbb{P}(W|D) = \frac{\mathbb{P}(W \cap D)}{\mathbb{P}(D)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

Therefore, the probability of winning increases.

**Example.** For the same game, you win if the sum of the outcomes is 7. The first toss is a 4. Should you be happy?

**Solution:** Considering the initial case, we have:

$$\mathbb{P}(W) = \frac{|W|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$$

Let  $D$  be the event that the first toss is a 4. Given  $D$ , we have:

$$\mathbb{P}(W|D) = \frac{\mathbb{P}(W \cap D)}{\mathbb{P}(D)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

Therefore, the probability does not change.

### 3.1.2 Properties of Conditional Probability

**Proposition 3.1.1 (Properties of Conditional Probability).** Conditional Probability  $\mathbb{P}(\cdot|D)$  are probabilities over reduced sample space  $F$  and satisfy probability axioms:

1. For every  $A$ ,  $\mathbb{P}(A|F) \geq 0$
2.  $\mathbb{P}(F|F) = 1$
3. For disjoint events  $A, B$ :  $\mathbb{P}(A \cup B|F) = \mathbb{P}(A|F) + \mathbb{P}(B|F)$

**Remark.** There is no assumption about  $A$  or  $B$  are subsets of  $F$ .

We can then generalize the conditional probability using Uniform Probability Law. Under equally-likely outcomes in  $F$ :

$$\mathbb{P}(A|F) = \frac{\text{Number of outcomes in } A \cap F}{\text{Number of outcomes in } F} = \frac{|A \cap F|}{|F|}$$

**Example.**

There are three cards with Black/Black, Red/Red and Red/Black sides. We randomly draw one card and observe that one of its sides is Black.



What is the probability that the card's other side is also Black?

**Solution:** Let  $B$  be the event that the first drawn color is black, and  $E$  be the event that the second drawn color is also black. Then we have:

$$\Omega = \{1F, 1B, 2F, 2B, 3F, 3B\}$$

$$\mathbb{P}(E|B) = \frac{|E \cap B|}{|B|} = \frac{2}{3}$$

### 3.1.3 The Multiplication Rule

**Proposition 3.1.2 (The Multiplication Rule).** For events  $E_1, E_2$  we can write the probability of their intersection  $E_1 \cap E_2$  as

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1)\mathbb{P}(E_2|E_1)$$

In general, for every  $E_1, E_2, \dots, E_n$ , the multiplication rule says

$$\mathbb{P}(E_1 \cap E_2 \cap \dots \cap E_n) = \mathbb{P}(E_1)\mathbb{P}(E_2|E_1)\mathbb{P}(E_3|E_1 \cap E_2) \dots \mathbb{P}(E_n|E_1 \cap E_2 \cap \dots \cap E_{n-1})$$

**Example.** A box contains 5 red balls and 15 blue balls. We randomly draw three balls from the box (without replacement). What is the probability that the balls are all red?

**Solution:** Let  $R_i$  be the event that the  $i$ -th ball being drawn is red.

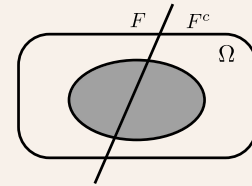
Then we have

$$\mathbb{P}(R_1 \cap R_2 \cap R_3) = \mathbb{P}(R_1)\mathbb{P}(R_2|R_1)\mathbb{P}(R_3|R_1 \cap R_2) = \frac{5}{20} \times \frac{4}{19} \times \frac{3}{18} = \frac{1}{114}$$

**Theorem 3.1.1 (Total Probability Theorem).**

For every event  $E$  and  $F$  and its complement  $F^c$ ,

$$\begin{aligned} \mathbb{P}(E) &= \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c) \\ &= \mathbb{P}(E|F)\mathbb{P}(F) + \mathbb{P}(E|F^c)\mathbb{P}(F^c) \end{aligned}$$



More generally, if events  $F_1, F_2, \dots, F_n$  partition  $\Omega$  (disjoint events and  $F_1 \cup F_2 \cup \dots \cup F_n = \Omega$ ), then the total probability theorem says

$$\mathbb{P}(E) = \mathbb{P}(E|F_1)\mathbb{P}(F_1) + \mathbb{P}(E|F_2)\mathbb{P}(F_2) + \dots + \mathbb{P}(E|F_n)\mathbb{P}(F_n)$$

**Example.** A box contains 5 red balls and 15 blue balls. We randomly draw two balls from the box (without replacement). What is the probability that the balls have different colors?

**Solution:** Let  $R$  be the event that the first ball is red,  $E$  be the event that the balls have different colors. Then we have

$$\mathbb{P}(E) = \mathbb{P}(E|R)\mathbb{P}(R) + \mathbb{P}(E|R^c)\mathbb{P}(R^c) = \frac{15}{19} \times \frac{5}{20} + \frac{5}{19} \times \frac{15}{20} = \frac{15}{38}$$

**Example.** For the situation that a group of students answered one multiple choice question where there are 4 options. What is the probability that a student knows the answer to the question?

**Solution:** Here we define event  $K$  as student knows the answer to the question, and event  $C$  as student correctly answers the question. Then, the total probability theorem says

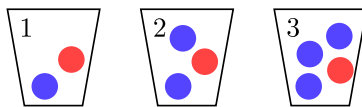
$$\begin{aligned} \mathbb{P}(C) &= \mathbb{P}(C|K)\mathbb{P}(K) + \mathbb{P}(C|K^c)\mathbb{P}(K^c) \\ &= 1 \times \mathbb{P}(K) + \frac{1}{4} \times (1 - \mathbb{P}(K)) \\ &= \frac{3}{4}\mathbb{P}(K) + \frac{1}{4} \end{aligned}$$

This gives

$$\mathbb{P}(K) = \frac{4\mathbb{P}(C) - 1}{3}$$

## 3.2 Bayes' Rule

We choose a cup at random and then a random ball from that cup. The selected ball is red. Which cup do you guess the ball came from?



This is an example of cause and effect. Here, the cause is the cup number and the effect is the ball's color. Choosing different cups leads to different probabilities of choosing the color of balls.

**Theorem 3.2.1 (Bayes' Rule).** Consider events  $C$  and  $E$ . Then,

$$\mathbb{P}(C|E) = \frac{\mathbb{P}(E|C)\mathbb{P}(C)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|C)\mathbb{P}(C)}{\mathbb{P}(E|C)\mathbb{P}(C) + \mathbb{P}(E|C^c)\mathbb{P}(C^c)}$$

More generally, if  $C_1, C_2, \dots, C_n$  partition the set of possible causes  $S$ ,

$$\mathbb{P}(C_1|E) = \frac{\mathbb{P}(E|C_1)\mathbb{P}(C_1)}{\mathbb{P}(E|C_1)\mathbb{P}(C_1) + \mathbb{P}(E|C_2)\mathbb{P}(C_2) + \dots + \mathbb{P}(E|C_n)\mathbb{P}(C_n)}$$

Let us revisit the example above. Let  $R$  represent the event that a red ball is drawn, and let  $C_i$  denote the cup from which the red ball is drawn, where  $i = 1, 2, 3$ . We can then calculate the probability of each cup being the source of the red ball, given that the ball is red.

$$\mathbb{P}(C_i|R) = \frac{\mathbb{P}(R|C_i)\mathbb{P}(C_i)}{\mathbb{P}(R|C_1)\mathbb{P}(C_1) + \mathbb{P}(R|C_2)\mathbb{P}(C_2) + \mathbb{P}(R|C_3)\mathbb{P}(C_3)}$$

For sample space, we have:  $\Omega = \{1B, 1R, 2B_1, 2B_2, 2R, 3B_1, 3B_2, 3B_3, 3R\}$  (not equally likely)

	Cup 1	Cup 2	Cup 3
$\mathbb{P}(C_i)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\mathbb{P}(R C_i)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$

For  $\mathbb{P}(R)$ , by using total probability theorem, we have

$$\mathbb{P}(R) = \mathbb{P}(R|C_1)\mathbb{P}(C_1) + \mathbb{P}(R|C_2)\mathbb{P}(C_2) + \mathbb{P}(R|C_3)\mathbb{P}(C_3) = \frac{1}{2} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{4} \times \frac{1}{3} = \frac{13}{36}$$

Then we have:

$$\begin{aligned} \mathbb{P}(C_1|R) &= \frac{\mathbb{P}(R|C_1)\mathbb{P}(C_1)}{\mathbb{P}(R)} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{13}{36}} = \frac{6}{13}, \quad \mathbb{P}(C_2|R) = \frac{\mathbb{P}(R|C_2)\mathbb{P}(C_2)}{\mathbb{P}(R)} = \frac{\frac{1}{3} \times \frac{1}{3}}{\frac{13}{36}} = \frac{4}{13} \\ \mathbb{P}(C_3|R) &= \frac{\mathbb{P}(R|C_3)\mathbb{P}(C_3)}{\mathbb{P}(R)} = \frac{\frac{1}{4} \times \frac{1}{3}}{\frac{13}{36}} = \frac{3}{13} \end{aligned}$$

**Example.** Two classes take place in the same academic building. Class A has 100 students from whom 20% are female, Class B has 10 students from whom 80% are female.

Now we see a female student in this building. What is the probability that the student is from Class A?

**Solution:**

$$\mathbb{P}(A|F) = \frac{\mathbb{P}(F|A)\mathbb{P}(A)}{\mathbb{P}(F|A)\mathbb{P}(A) + \mathbb{P}(F|B)\mathbb{P}(B)} = \frac{20\% \times \frac{10}{11}}{20\% \times \frac{10}{11} + 80\% \times \frac{1}{11}} = \frac{5}{7}$$

This example demonstrates a counterintuitive result. Although the proportion of females in Class B is greater than that in Class A, the probability of a female student belonging to Class A is higher. This highlights the importance of using Bayes' rule to update our beliefs, as our initial hypotheses may not always align with the actual probabilities.

To summarize, we see that conditional probability will be a very powerful tool when

1. the studied environment includes causes and effect. This is especially useful for calculating the probability of a cause under an observation of the effect.
2. we want to calculate an ordinary probability and conditioning on the right event can simplify the description of the sample space.

## 3.3 Independence

### 3.3.1 Independence of Two Events

**Definition 3.3.1 (Independent Events).** We call Events  $A$  and  $B$  independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \quad \text{or equivalently} \quad \mathbb{P}(A|B) = \mathbb{P}(A)$$

**Example.** We toss a coin three times. Then the following events are independent:

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

Event  $E_1$ : The first toss show heads

Event  $E_2$ : The second and third toss both show tails

$$\text{We have: } \mathbb{P}(E_1) = \frac{4}{8}, \mathbb{P}(E_2) = \frac{2}{8}, \mathbb{P}(E_1 \cap E_2) = \frac{1}{8}, \mathbb{P}(E_1)\mathbb{P}(E_2) = \frac{4}{8} \times \frac{2}{8} = \frac{1}{8} = \mathbb{P}(E_1 \cap E_2)$$

This shows that the events are independent.

**Example.** We roll two dice. Consider events:

$E_1$ : the first die is 4;  $S_6$ : the sum of dice is 6;  $S_7$ : the sum of dice is 7.

$$\mathbb{P}(E_1) = \frac{1}{6}; \mathbb{P}(S_6) = \frac{5}{36}; \mathbb{P}(S_7) = \frac{1}{6}; \mathbb{P}(E_1 \cap S_6) = \frac{1}{36}; \mathbb{P}(E_1 \cap S_7) = \frac{1}{36}; \mathbb{P}(S_6 \cap S_7) = 0.$$

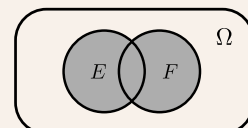
Then we know that  $(E_1, S_6)$  and  $(S_6, S_7)$  are not independent,  $(E_1, S_7)$  are independent.

**Remark.** Mutually exclusive does not mean independent.

#### Proposition 3.3.1.

If  $E, F$  are independent events, then events  $E^c, F$  will also be independent.

$$\mathbb{P}(E^c \cap F) = \mathbb{P}(E^c)\mathbb{P}(F)$$



**Proof.**

$$\begin{aligned} \mathbb{P}(E^c \cap F) &= \mathbb{P}(F) - \mathbb{P}(E \cap F) \\ &= \mathbb{P}(F) - \mathbb{P}(E)\mathbb{P}(F) \\ &= \mathbb{P}(F)(1 - \mathbb{P}(E)) \\ &= \mathbb{P}(E^c)\mathbb{P}(F) \end{aligned}$$

■

### 3.3.2 Independence of Several Events

We call three events  $A, B, C$  independent events if these four conditions are satisfied:

1.  $A$  and  $B$  are independent:  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
2.  $B$  and  $C$  are independent:  $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$
3.  $A$  and  $C$  are independent:  $\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$
4. And we require  $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$

**Example.** We roll two dice. Consider events:

$E_1$ : the first die is 4;  $E_2$ : the second die is 3;  $S_7$ : the sum of dice is 7.

$$\mathbb{P}(E_1) = \frac{1}{6}; \mathbb{P}(E_2) = \frac{1}{6}; \mathbb{P}(S_7) = \frac{1}{6}; \mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1 \cap S_7) = \mathbb{P}(E_2 \cap S_7) = \frac{1}{36}.$$

Then we know that  $(E_1, E_2)$ ,  $(E_1, S_7)$  and  $(E_2, S_7)$  are independent.

$$\mathbb{P}(E_1 \cap E_2 \cap S_7) = \frac{1}{36} \neq \mathbb{P}(E_1) \times \mathbb{P}(E_2) \times \mathbb{P}(S_7)$$

Then we know that  $(E_1, E_2, S_7)$  are not independent.

We call  $n$  events  $A_1, A_2, \dots, A_n$  independent events if for every subset  $\{j_1, j_2, \dots, j_t\}$  of  $\{A_1, \dots, A_n\}$ , the probability of the intersection is the product of their probabilities:

$$\mathbb{P}(A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_t}) = \mathbb{P}(A_{j_1})\mathbb{P}(A_{j_2}) \dots \mathbb{P}(A_{j_t})$$

If  $n$  events  $A_1, A_2, \dots, A_n$  are independent events, then the independence is preserved when we replace some event(s) by their complements, intersection, unions.

For example, if  $A, B, C, D$  are independent events, then  $A \cup B$  and  $C \cap D$  are also independent events.

**Example.** Alice wins 60% of her ping pong matches against Bob. They meet for a 3 match playoff. What are the chances that Alice will win the playoff?

**Solution** For Alice, we have

$$\Omega = \{\text{WWW}, \text{WWL}, \text{WLW}, \text{LWW}, \text{WLL}, \text{LWL}, \text{LLW}, \text{LLL}\}$$

$$A = \{\text{WWW}, \text{WWL}, \text{WLW}, \text{LWW}\}$$

$$\mathbb{P}(A) = 0.6^3 + 0.6^2 * 0.4 * 3 = \frac{81}{125}$$

### 3.3.3 Conditional Independence

**Definition 3.3.2 (Conditional Independence).** Events  $A$  and  $B$  are independent conditioned on event  $F$  if

$$\mathbb{P}(A \cap B | F) = \mathbb{P}(A | F)\mathbb{P}(B | F)$$

Note that the above is equivalent to

$$\mathbb{P}(A | B \cap F) = \mathbb{P}(A | F)$$



---

**Example.**

Today	Tomorrow
Sunny	80% Sunny, 20% Rainy
Rainy	40% Sunny, 60% Rainy

If Today(Monday) is rainy, what is the probability that Wednesday will also be sunny?

**Solution:** We suppose weather on Monday and Wednesday are independent conditioned on weather on Tuesday. Let  $M$  be the event that Monday is sunny, same for  $T$  and  $W$ . Then:

$$\begin{aligned}\mathbb{P}(W|M^c) &= \mathbb{P}(W|M^c \cap T)\mathbb{P}(T|M^c) + \mathbb{P}(W|M^c \cap T^c)\mathbb{P}(T^c|M^c) \\ &= 80\% \times 40\% + 40\% \times 60\% \\ &= 0.56\end{aligned}$$

## Chapter 4

# Random Variables

### 4.1 Introduction

**Definition 4.1.1 (Discrete Random Variable).** A discrete random variable assigns a discrete value to every outcome in the sample space  $\Omega$ .

For example, we can assign the number of heads in tossing 3 coins to the random variable  $X$ . Then we have:

$$\mathbb{P}(X = 0) = \frac{1}{8}, \mathbb{P}(X = 1) = \frac{3}{8}, \mathbb{P}(X = 2) = \frac{3}{8}, \mathbb{P}(X = 3) = \frac{1}{8}.$$

#### 4.1.1 Probability Mass Function (PMF)

**Definition 4.1.2 (Probability Mass Function).** The Probability Mass Function (PMF)  $p : \mathbb{R} \rightarrow [0, 1]$  of a discrete random variable  $X$  is the function

$$p(x) = \mathbb{P}(X = x)$$

We can describe the PMF by a table:

$x$	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

For every random variable  $X$ , its probability mass function satisfies the following (based on the axioms of probability):

1. For every  $x \in \mathbb{R}$ ,  $p(x) \geq 0$  is non-negative
2. If  $\mathcal{X}$  is the set of all possible values of  $X$ , then the PMF values on  $\mathcal{X}$  will add up to 1:

$$\sum_{x \in \mathcal{X}} p(x) = 1.$$

**Example.** We roll two 3-sided dice. Let random variable  $D$  be the difference between the output of the first and second dice. What is the PMF of random variable  $D$ ? What is the probability that  $D \geq 1$ ?

**Solution:** For sample space, we have

$$\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (3, 3)\} \implies D = x, x \in [-2, +2]$$

For PMF, we have:

$x$	$-2$	$-1$	$0$	$1$	$2$
$p(x)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

Then, for the probability that  $D \geq 1$ , we have

$$\mathbb{P}(D \geq 1) = p(1) + p(2) = \frac{2}{9} + \frac{1}{9} = \frac{3}{9}$$

## 4.2 Binomial Random Variable

### 4.2.1 Definition

We call  $X$  a Binomial  $(n, p)$  Random Variable when  $X$  represents the number of successes over  $n$  independent trials, each with a success probability of  $p$ .

For example, if we toss  $n$  coins, the number of heads is  $\text{Binomial}(n, \frac{1}{2})$

**Example.** We flip a coin 10 times and consider the random variable of the number of consecutive changes (HT or TH) in the 10 coin flips. What is this random variable?

**Solution:** We have a total of 9 trials, with the possible outcomes being HH, HT, TH, and TT. The probability of a consecutive change is  $\frac{1}{2}$ . This gives  $\text{Binomial}(9, \frac{1}{2})$ .

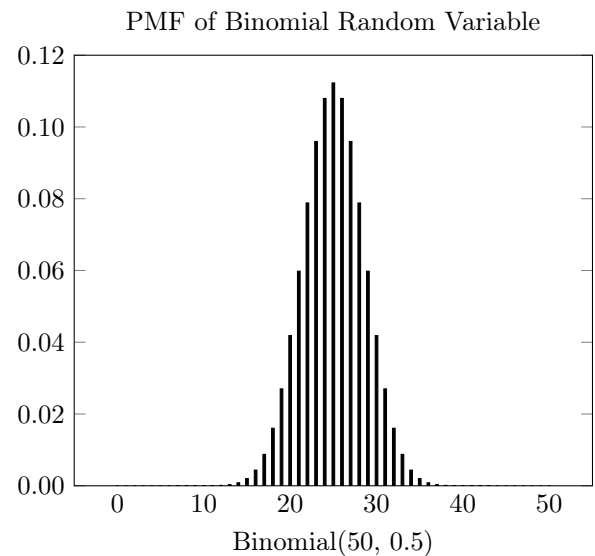
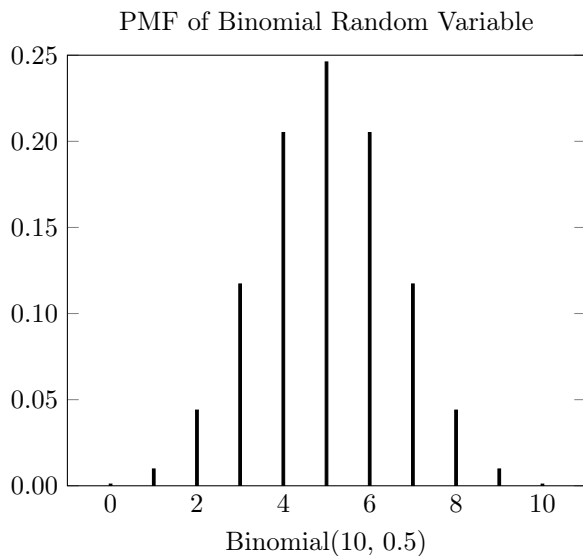
**Example.** We draw a 10-card hand from a 52-card deck. Let  $N$  = Number of Aces among the picked cards. What is the random variable  $N$ ?

**Solution:** The definition states that the events must be independent. However, the second draw is influenced by the outcome of the first draw, so the events are not independent.

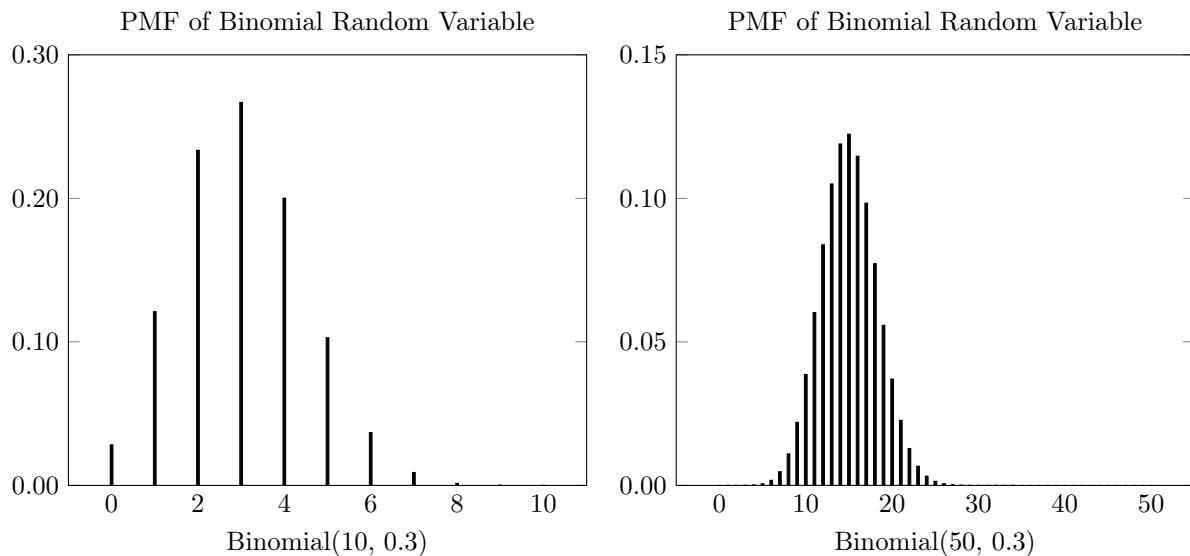
### 4.2.2 PMF of Binomial Random Variable

The probability mass function (PMF) of a  $\text{Binomial}(n, p)$  Random Variable is

$$p(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$



From the chart above, we can observe that as the value of  $n$  increases, the graph is symmetric around the central point. For  $\text{Binomial}(10, 0.5)$ , the probability reaches its maximum at  $x = 5$ , which can also be calculated as  $10 \times 0.5$ . The same pattern holds for the graph on the right-hand side. This also makes sense because the probability of the event occurring is 0.5. Therefore, we can estimate that half of the events will occur, resulting in the maximum probability at  $x = 5$ .



For the two graphs above, it also holds true that the maximum value is reached at  $10 \times 0.3 = 3$  or  $50 \times 0.3 = 15$ , due to the same reasoning. However, in this case, the graph is no longer symmetric around the central point.

Additionally, we observe that the graph shifts to the left when  $p < 0.5$ . Conversely, if  $p > 0.5$ , the graph shifts to the right.

**Example.** The Lakers and the Celtics meet for a 7-game playoff.

Lakers win 60% of the time. What is the probability that all 7 games are played? What is the probability that Lakers win the play-off in 6 games?

**Solution:** For all 7 games to be played, the first 6 games must occur, which means that both the Lakers and the Celtics win 3 games each. So  $X$  is a  $\text{Binomial}(6, 0.6)$ .

$$\mathbb{P}(X = 3) = \binom{6}{3} \times (0.6)^3 \times (1 - 0.6)^3 = \frac{864}{3125}$$

For the Lakers to win in 6 games, we cannot simply use  $\mathbb{P}(X = 4)$  because it also includes the probability of the Lakers winning in 4 or 5 games. However, by fixing the Lakers to win the 6th game, we can define a new variable following  $\text{Binomial}(5, 0.6)$ . We then only need to calculate the probability of the Lakers winning 3 games in the first 5 games.

$$\mathbb{P}(X = 3) \times 0.6 = \binom{5}{3} \times (0.6)^3 \times (1 - 0.6)^2 \times 0.6 = \frac{648}{3125}$$

## 4.3 Geometric Random Variable

### 4.3.1 Definition

We call  $N$  a  $\text{Geometric}(p)$  Random Variable when  $X$  represents the first time of success over a series of independent trials  $X_1, X_2, \dots$ , each with a success probability of  $p$ :

$$N = \text{first (smallest) } n \text{ such that } X_n = 1.$$

For example, we toss a coin until we see the first heads. The number of coin tosses to see the first heads is  $\text{Geometric}(\frac{1}{2})$ .

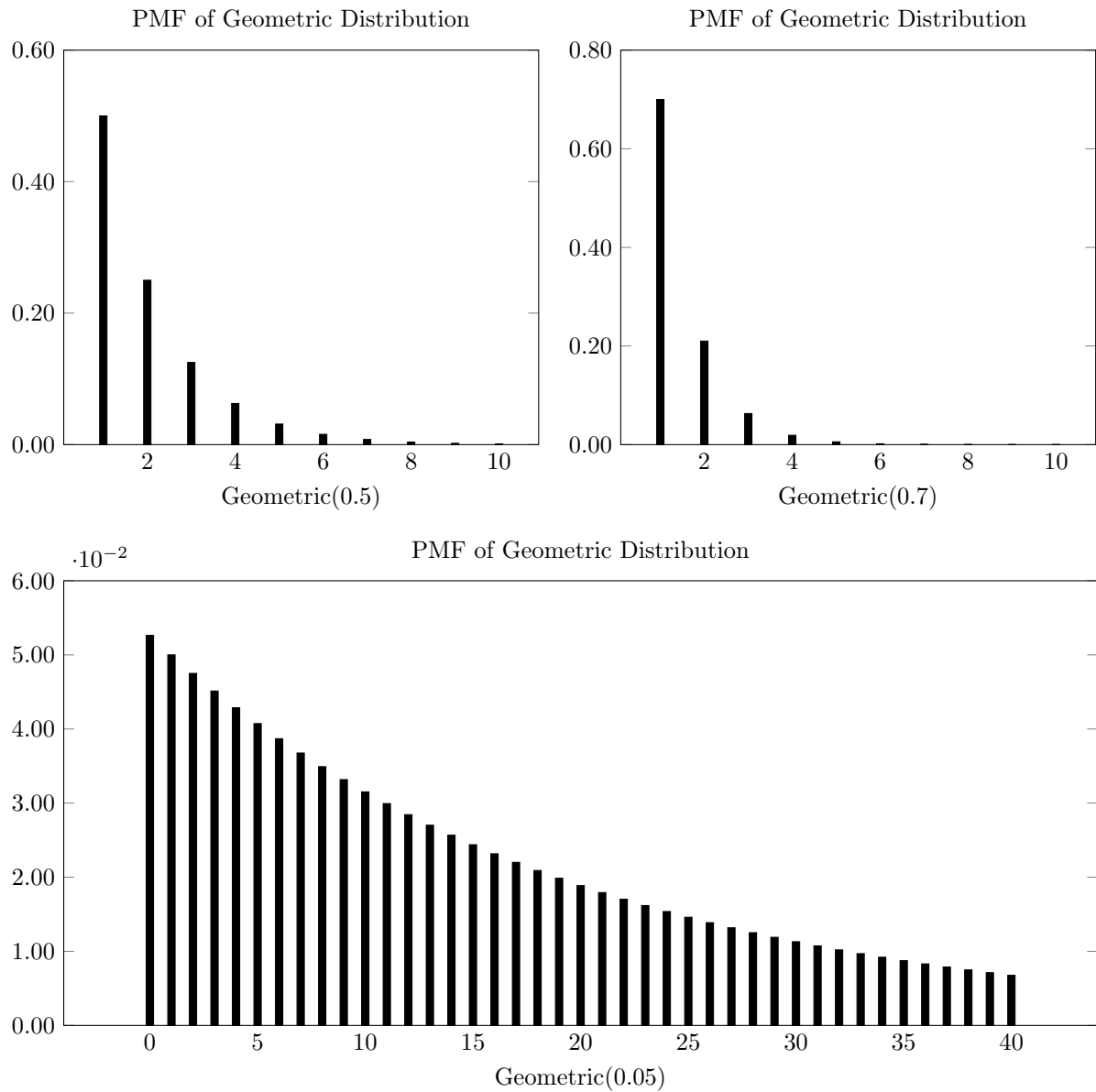
### 4.3.2 PMF of Geometric Random Variable

The probability mass function (PMF) of a  $\text{Geometric}(p)$  Random Variable is

$$p(k) = \mathbb{P}(X = k) = p(1 - p)^{k-1}$$

Axioms of probability also holds for PMF of geometric random variable. Therefore,

$$\sum_{k=1}^{\infty} p(1 - p)^{k-1} = 1$$



**Example.** You keep rolling dice until you roll a 6. What is the probability that you rolled more than 10 times?

**Solution:** Define  $X$  is a Geometric( $\frac{1}{6}$ ), then we have

$$\mathbb{P}(X \geq 11) = \sum_{k=11}^{\infty} p(1-p)^{k-1} = \sum_{k=11}^{\infty} \frac{1}{6} \times \left(\frac{5}{6}\right)^{k-1} = \left(\frac{5}{6}\right)^{10}$$

By generalizing the above, we have

$$\sum_{k=m}^{\infty} p(1-p)^{k-1} = (1-p)^{m-1}$$

This formula means that if we want to find the probability of success occurring at or after the  $m$ -th trial, we can treat the trials before the  $m$ -th one as all failures. This simplifies the calculation.

## 4.4 Cumulative Distribution Function (CDF)

**Definition 4.4.1 (Cumulative Distribution Function (CDF)).** For a random variable  $X$ , its cumulative distribution function (CDF)  $F(x)$  is:

$$F(x) = \mathbb{P}(X \leq x)$$

From the definition of PMF, we have

$$F(x) = \sum_{\substack{k \in \mathcal{X} \\ k \leq x}} p(x)$$

For a Geometric( $p$ ) random variable  $X$ , the CDF will be

$$F(k) = \mathbb{P}(X \leq k) = 1 - \mathbb{P}(X > k) = 1 - (1-p)^k$$

For a Binomial( $n, p$ ) random variable  $X$ , the CDF will be

$$F(k) = \mathbb{P}(X \leq k) = (1-p)^n + n(1-p)^{n-1}p + \dots + \binom{n}{k}(1-p)^{n-k}p^k$$

**Example.** You keep rolling dice until you roll a 6. What is the probability that you roll the dice an even number of times?

**Solution 1:**

$$\begin{aligned} \mathbb{P}(X \text{ is even}) &= \mathbb{P}(X = 2) + \mathbb{P}(X = 4) + \mathbb{P}(X = 6) + \dots \\ &= p(2) + p(4) + p(6) + \dots \\ &= \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + \frac{1}{6} \times \left(\frac{5}{6}\right)^5 + \dots \end{aligned}$$

**Solution 2:** Let's define event  $A$  as the event that an even number is rolled. Next, we define another event  $B$ , which is independent of event  $A$ , such that using conditional probability, we have  $\mathbb{P}(A) = \mathbb{P}(A|B)$ .

For example, we can define  $B = \{X = 1 \text{ or } X = 2\}$ . Since  $A$  and  $B$  are independent, we can then proceed with the calculation.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(X = 2)}{\mathbb{P}(X = 1) + \mathbb{P}(X = 2)} = \frac{p(1-p)}{p + p(1-p)} = \frac{p-1}{p-2} = \frac{\frac{5}{6}}{\frac{11}{6}} = \frac{5}{11}$$

## 4.5 Poisson Random Variable

**Example.** Alice randomly sprinkles 25 chocolate chips on 5 cookies.

1. What is the random variable  $N$  on how many chips a cookie gets?

$$N \sim \text{Binomial}(25, \frac{1}{5})$$

2. What is the probability a cookie gets no chips?

$$\mathbb{P}(N = 0) = \binom{25}{0} \times \left(\frac{1}{5}\right)^0 \times \left(1 - \frac{1}{5}\right)^{25} = 0.004$$

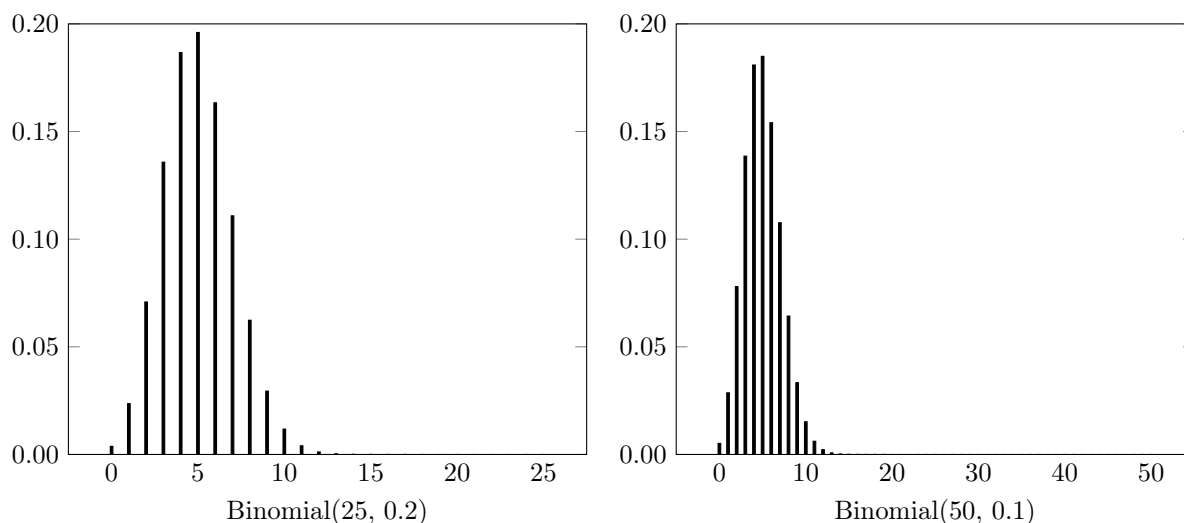
3. What is the probability a cookie gets exactly 5 chips?

$$\mathbb{P}(N = 5) = \binom{25}{5} \times \left(\frac{1}{5}\right)^5 \times \left(1 - \frac{1}{5}\right)^{20} = 0.196$$

4. What is the probability a cookie gets at most 5 chips?

$$\mathbb{P}(N \leq 5) \approx 0.617$$

We now want to examine how the probability changes if we increase the values, for example, by having 250 chocolate chips and 50 cookies instead. Intuitively, the probability should not change significantly, and through calculation, the results are indeed similar to those obtained in the previous example.



From the above, we see that when we double the value, the PMF remains quite similar to the previous one. Additionally, the average rate of chocolate chips per cookie is also the same, i.e.,  $25 \times 0.2 = 50 \times 0.1 = 5$ .

Poisson(5)	$\mathbb{P}(X = 0)$	$\mathbb{P}(X = 5)$	$\mathbb{P}(X \leq 5)$
Binomial(25, 0.2)	0.004	0.196	0.617
Binomial(50, 0.1)	0.005	0.185	0.616
Binomial(500, 0.01)	0.006	0.176	0.615

From the values above, we can see that the probability converges to a certain value. This is what we call a Poisson Random Variable. The values we obtain by using  $25 \times 0.2$ ,  $50 \times 0.1$  represent the rate of Poisson random variable.

**Definition 4.5.1 (Poisson Random Variable).** A  $\text{Poisson}(\lambda)$  random variable  $X$  has the PMF:

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, 3, \dots$$

where  $\lambda$  is called the rate parameter.

**Theorem 4.5.1.**  $\text{Poisson}(\lambda)$  is the limit approximation of  $\text{Binomial}(n, \frac{\lambda}{n})$  when  $n \rightarrow \infty$ :

$$\mathbb{P}(\text{Poisson}(\lambda) = k) = \lim_{n \rightarrow \infty} \mathbb{P}(\text{Binomial}(n, \frac{\lambda}{n}) = k)$$

Poisson random variable can be used when we have a large number of independent trials while the expected number of successes remains small (the rate of successes being constant).

**Example.** Suppose rain is falling on your head a rate of 3 drops/sec. What is the probability that you get

(1) no hits in the next second?

**Solution:**

$$\mathbb{P}(\text{Poisson}(3) = 0) = e^{-3} \frac{3^0}{0!} = e^{-3} \approx 0.050$$

(2) at least 3 hits in the next second?

**Solution:**

$$\mathbb{P}(\text{Poisson}(3) \geq 3) = 1 - \mathbb{P}(\text{Poisson}(3) < 3) = 1 - e^{-3} \frac{3^0}{0!} - e^{-3} \frac{3^1}{1!} - e^{-3} \frac{3^2}{2!} = 1 - \frac{17}{2e^3} \approx 0.576$$

(3) exactly 10 hits in the next 5 seconds?

**Solution:**

$$\mathbb{P}(\text{Poisson}(3 \times 5) = 10) = e^{-15} \frac{15^{10}}{10!} = 0.049$$

## 4.6 Properties of Random Variables

### 4.6.1 Expected Value

**Definition 4.6.1 (Expected Value).** The expected value (expectation) of a random variable  $X$  with PMF  $p(x)$  is

$$\mathbb{E}[X] = \sum_x xp(x)$$

For example, the expected value of random variable  $X$  as the number of heads in tossing 1 coin:

$$\mathbb{E}[X] = \frac{1}{2} \times 0 + \frac{1}{2} \times 1 = \frac{1}{2}$$

Now, instead of flipping one coins, we flip 3 coins. For PMF, we have

$$p(k) = \binom{3}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{3-k} = \frac{\binom{3}{k}}{8}$$

Then, the expected value of random variable  $X$  as the number of heads in tossing 3 coin:

$$\mathbb{E}[X] = \frac{1}{8} \times 0 + \frac{3}{8} \times 1 + \frac{3}{8} \times 2 + \frac{1}{8} \times 3 = \frac{3}{2}$$

The expectation is the average value the random variable takes when the experiment is done many times.



**Example.** Find the expected value of random variable  $F$  as the face value of a six-sided die.

**Solution:**

$$\mathbb{E}[X] = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2}$$

**Example.** We play this game that we roll three dice and

- You win  $k$  dollars if we see  $k \geq 1$  "Two" outcomes.
- You lose 1 dollar if we see no "Two" outcomes.

Should you play this game?

**Solution:**

$$G = -1, \mathbb{P}(-1) = \left(\frac{5}{6}\right)^3 = \frac{125}{216}; \quad G = 1, \mathbb{P}(1) = \binom{3}{1} \left(\frac{1}{6}\right) \left(\frac{5}{6}\right)^2 = \frac{25}{72};$$

$$G = 2, \mathbb{P}(2) = \binom{3}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^1 = \frac{5}{72}; \quad G = 3, \mathbb{P}(3) = \left(\frac{1}{6}\right)^3 = \frac{1}{216};$$

$$\mathbb{E}[X] = -1 \times \frac{125}{216} + 1 \times \frac{25}{72} + 2 \times \frac{5}{72} + 3 \times \frac{1}{216} = -\frac{17}{216} \approx -0.079$$

Therefore, it's better not to play this game.

#### 4.6.2 Function of Random Variables

If  $X$  is a random variable with PMF  $p_X$ , then  $Y = f(X)$  will also be a random variable with PMF  $p_Y$ :

$$p_Y(y) = \sum_{x: f(x)=y} p_X(x)$$

The expected value of  $f(X)$  for a function  $f$  and random variable  $X$  is

$$\mathbb{E}[f(X)] = \sum_x f(x)p(x)$$

**Example.** Consider a random variable  $X$  with the following PMF:

$x$	0	1	2
$p(x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

$y = X - 1$	-1	0	1
$p(y)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

$z = (X - 1)^2$	0	1
$p(z)$	$\frac{1}{3}$	$\frac{2}{3}$

Using  $Y = (X - 1)^2$  as an example, we have:

$$p_Y(1) = \sum_{x: (x-1)^2=1} p_X(x) = p_X(0) + p_X(2) = \frac{2}{3}$$

Additionally, for the expected values of the above PMF, we have:

$$\begin{aligned} \mathbb{E}[X] &= 0 \times \frac{1}{3} + 1 \times \frac{1}{3} + 2 \times \frac{1}{3} = 1 \\ \mathbb{E}[X - 1] &= -1 \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = 0 \\ \mathbb{E}[(X - 1)^2] &= 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3} \end{aligned}$$

Alternatively, we can use the following method to directly calculate the expected values:

$$\begin{aligned}\mathbb{E}[X - 1] &= \sum_{x \in \{0,1,2\}} (x - 1)p(x) = (0 - 1) \times \frac{1}{3} + (1 - 1) \times \frac{1}{3} + (2 - 1) \times \frac{1}{3} = 0 \\ \mathbb{E}[(X - 1)^2] &= \sum_{x \in \{0,1,2\}} (x - 1)^2 p(x) = (0 - 1)^2 \times \frac{1}{3} + (1 - 1)^2 \times \frac{1}{3} + (2 - 1)^2 \times \frac{1}{3} = \frac{2}{3}\end{aligned}$$

**Remark.**

$$\mathbb{E}[f(x)] \neq f(\mathbb{E}[X])$$

**Example.** Suppose the distance between place A and place B is 1 km. There is a 60% chance that the weather will be sunny; in that case, you will walk from A to B at a speed of 5 km/h. Conversely, there is a 40% chance that the weather will be rainy; in that case, you will take a shuttle traveling at a speed of 30 km/h. Find the expected value of time  $T$  and speed  $V$ .

**Solution:** Given that  $V = \frac{1}{T}$ ,

$$\begin{aligned}\mathbb{E}[V] &= 0.6 \times 5 + 0.4 \times 30 = 15 \text{ km/h} \\ \mathbb{E}[T] &= \mathbb{E}\left[\frac{1}{V}\right] = 0.6 \times \frac{1}{5} + 0.4 \times \frac{1}{30} = \frac{2}{15} \text{ h} \neq \frac{1}{\mathbb{E}[V]}\end{aligned}$$

## Chapter 5

### 5.1

## Chapter 6

### 6.1

## Chapter 7

### 7.1

## Chapter 8

### 8.1

## Chapter 9

### 9.1

## Chapter 10

### 10.1