

# Chapter 4 Probability Theory

Chances are that you will face at least a couple of probability problems in most quantitative interviews. Probability theory is the foundation of every aspect of quantitative finance. As a result, it has become a popular topic in quantitative interviews.

Although good intuition and logic can help you solve many of the probability problems, having a thorough understanding of basic probability theory will provide you with clear and concise solutions to most of the problems you are likely to encounter. Furthermore, probability theory is extremely valuable in explaining some of the seemingly-counterintuitive results. Armed with a little knowledge, you will find that many of the interview problems are no more than disguised textbook problems.

So we dedicate this chapter to reviewing basic probability theory that is not only broadly tested in interviews but also likely to be helpful for your future career.<sup>1</sup> The knowledge is applied to real interview problems to demonstrate the power of probability theory. Nevertheless, the necessity of knowledge in no way downplays the role of intuition and logic. Quite the contrary, common sense and sound judgment are always crucial for analyzing and solving either interview or real-life problems. As you will see in the following sections, all the techniques we discussed in Chapter 2 still play a vital role in solving many of the probability problems.

Let's have some fun playing the odds.

## 4.1 Basic Probability Definitions and Set Operations

First let's begin with some basic definitions and notations used in probability. These definitions and notations may seem dry without examples—which we will present momentarily—yet they are crucial to our understanding of probability theory. In addition, it will lay a solid ground for us to systematically approach probability problems.

**Outcome ( $\omega$ ):** the outcome of an experiment or trial.

**Sample space/Probability space ( $\Omega$ ):** the set of all possible outcomes of an experiment.

---

<sup>1</sup> As I have emphasized in Chapter 3, this book does not teach probability or any other math topics due to the space limit—it is not my goal to do so, either. The book gives a summary of the frequently-tested knowledge and shows how it can be applied to a wide range of real interview problems. The knowledge used in this chapter is covered by most introductory probability books. It is always helpful to pick up one or two classic probability books in case you want to refresh your memory on some of the topics. My personal favorites are *First Course in Probability* by Sheldon Ross and *Introduction to Probability* by Dimitri P. Bertsekas and John N. Tsitsiklis.

$P(\omega)$ : Probability of an outcome ( $P(\omega) \geq 0, \forall \omega \in \Omega, \sum_{\omega \in \Omega} P(\omega) = 1$ ).

**Event:** A set of outcomes and a subset of the sample space.

$P(A)$ : Probability of an event  $A$ ,  $P(A) = \sum_{\omega \in A} P(\omega)$ .

$A \cup B$ : Union  $A \cup B$  is the set of outcomes in event  $A$  or in event  $B$  (or both).

$A \cap B$  or  $AB$ : Intersection  $A \cap B$  (or  $AB$ ) is the set of outcomes in both  $A$  and  $B$ .

$A^c$ : The complement of  $A$ , which is the event “not  $A$ ”.

**Mutually Exclusive:**  $A \cap B = \Phi$  where  $\Phi$  is an empty set.

For any mutually exclusive events  $E_1, E_2, \dots, E_N$ ,  $P\left(\bigcup_{i=1}^N E_i\right) = \sum_{i=1}^N P(E_i)$ .

**Random variable:** A **function** that maps each outcome ( $\omega$ ) in the sample space ( $\Omega$ ) into the set of real numbers.

Let’s use the rolling of a six-sided dice to explain these definitions and notations. A roll of a dice has 6 possible outcomes (mapped to a random variable): 1, 2, 3, 4, 5, or 6. So the sample space  $\Omega$  is  $\{1, 2, 3, 4, 5, 6\}$  and the probability of each outcome is  $1/6$  (assuming a fair dice). We can define an event  $A$  representing the event that the outcome is an odd number  $A = \{1, 3, 5\}$ , then the complement of  $A$  is  $A^c = \{2, 4, 6\}$ . Clearly  $P(A) = P(1) + P(3) + P(5) = 1/2$ . Let  $B$  be the event that the outcome is larger than 3:  $B = \{4, 5, 6\}$ . Then the union is  $A \cup B = \{1, 3, 4, 5, 6\}$  and the intersection is  $A \cap B = \{5\}$ . One popular random variable called indicator variable (a binary dummy variable) for event  $A$  is defined as the following:

$I_A = \begin{cases} 1, & \text{if } x \in \{1, 3, 5\} \\ 0, & \text{if } x \notin \{1, 3, 5\} \end{cases}$ . Basically  $I_A = 1$  when  $A$  occurs and  $I_A = 0$  if  $A^c$  occurs. The expected value of  $I_A$  is  $E[I_A] = P(A)$ .

Now, time for some examples.

## Coin toss game

Two gamblers are playing a coin toss game. Gambler  $A$  has  $(n+1)$  fair coins;  $B$  has  $n$  fair coins. What is the probability that  $A$  will have more heads than  $B$  if both flip all their coins?<sup>2</sup>

*Solution:* We have yet to cover all the powerful tools probability theory offers. What do we have now? Outcomes, events, event probabilities, and surely our reasoning capabilities! The one extra coin makes  $A$  different from  $B$ . If we remove a coin from  $A$ ,  $A$  and  $B$  will become symmetric. Not surprisingly, the symmetry will give us a lot of nice properties. So let's remove the last coin of  $A$  and compare the number of heads in  $A$ 's first  $n$  coins with  $B$ 's  $n$  coins. There are three possible outcomes:

$E_1$ :  $A$ 's  $n$  coins have more heads than  $B$ 's  $n$  coins;

$E_2$ :  $A$ 's  $n$  coins have equal number of heads as  $B$ 's  $n$  coins;

$E_3$ :  $A$ 's  $n$  coins have fewer heads than  $B$ 's  $n$  coins.

By symmetry, the probability that  $A$  has more heads is equal to the probability that  $B$  has more heads. So we have  $P(E_1) = P(E_3)$ . Let's denote  $P(E_1) = P(E_3) = x$  and  $P(E_2) = y$ .

Since  $\sum_{\omega \in \Omega} P(\omega) = 1$ , we have  $2x + y = 1$ . For event  $E_1$ ,  $A$  will always have more heads

than  $B$  no matter what  $A$ 's  $(n+1)$ th coin's side is; for event  $E_3$ ,  $A$  will have no more heads than  $B$  no matter what  $A$ 's  $(n+1)$ th coin's side is. For event  $E_2$ ,  $A$ 's  $(n+1)$ th coin does make a difference. If it's a head, which happens with probability 0.5, it will make  $A$  have more heads than  $B$ . So the  $(n+1)$ th coin increases the probability that  $A$  has more heads than  $B$  by  $0.5y$  and the total probability that  $A$  has more heads is  $x + 0.5y = x + 0.5(1 - 2x) = 0.5$  when  $A$  has  $(n+1)$  coins.

## Card game

A casino offers a simple card game. There are 52 cards in a deck with 4 cards for each value 2, 3, 4, 5, 6, 7, 8, 9, 10, <sup>jack queen king ace</sup>  $J$ ,  $Q$ ,  $K$ ,  $A$ . Each time the cards are thoroughly shuffled (so each card has equal probability of being selected). You pick up a card from the deck and the dealer picks another one without replacement. If you have a larger number, you win; if the numbers are equal or yours is smaller, the house wins—as in all other casinos, the house always has better odds of winning. What is your probability of winning?

---

<sup>2</sup> Hint: What are the possible results (events) if we compare the number of heads in  $A$ 's first  $n$  coins with  $B$ 's  $n$  coins? By making the number of coins equal, we can take advantage of symmetry. For each event, what will happen if  $A$ 's last coin is a head? Or a tail?

*Solution:* One answer to this problem is to consider all 13 different outcomes of your card. The card can have a value 2, 3, ...,  $A$  and each has  $1/13$  of probability. With a value of 2, the probability of winning is  $0/51$ ; with a value of 3, the probability of winning is  $4/51$  (when the dealer picks a 2); ...; with a value of  $A$ , the probability of winning is  $48/51$  (when the dealer picks a 2, 3, ..., or  $K$ ). So your probability of winning is

$$\frac{1}{13} \times \left( \frac{0}{51} + \frac{4}{51} + \dots + \frac{48}{51} \right) = \frac{4}{13 \times 51} \times (0 + 1 + \dots + 12) = \frac{4}{13 \times 51} \times \frac{12 \times 13}{2} = \frac{8}{17}.$$

Although this is a straightforward solution and it elegantly uses the sum of an integer sequence, it is not the most efficient way to solve the problem. If you have got the core spirits of the coin tossing problem, you may approach the problem by considering three different outcomes:

$E_1$ : Your card has a number larger than the dealer's;

$E_2$ : Your card has a number equal to the dealer's;

$E_3$ : Your card has a number lower than the dealer's.

Again by symmetry,  $P(E_1) = P(E_3)$ . So we only need to figure out  $P(E_2)$ , the probability that two cards have equal value. Let's say you have randomly selected a card. Among the remaining 51 cards, only 3 cards will have the same value as your card. So the probability that the two cards have equal value is  $3/51$ . As a result, the probability that you win is  $P(E_1) = (1 - P(E_2))/2 = (1 - 3/51)/2 = 8/17$ .

## Drunk passenger

A line of 100 airline passengers are waiting to board a plane. They each hold a ticket to one of the 100 seats on that flight. For convenience, let's say that the  $n$ -th passenger in line has a ticket for the seat number  $n$ . Being drunk, the first person in line picks a random seat (equally likely for each seat). All of the other passengers are sober, and will go to their proper seats unless it is already occupied; In that case, they will randomly choose a free seat. You're person number 100. What is the probability that you end up in your seat (i.e., seat #100) ?<sup>3</sup>

*Solution:* Let's consider seats #1 and #100. There are two possible outcomes:

---

<sup>3</sup> Hint: If you are trying to use complicated conditional probability to solve the problem, go back and think again. If you decide to start with a simpler version of the problem, starting with two passengers and increasing the number of passengers to show a pattern by induction, you can solve the problem more efficiently. But the problem is much simpler than that. Focus on events and symmetry and you will have an intuitive answer.

$E_1$ : Seat #1 is taken before #100;

$E_2$ : Seat #100 is taken before #1.

If any passenger takes seat #100 before #1 is taken, surely you will not end up in your own seat. But if any passenger takes #1 before #100 is taken, you will definitely end up in your own seat. By symmetry, either outcome has a probability of 0.5. So the probability that you end up in your seat is 50%.

In case this over-simplified version of reasoning is not clear to you, consider the following detailed explanation: If the drunk passenger takes #1 by chance, then it's clear all the rest of the passengers will have the correct seats. If he takes #100, then you will not get your seat. The probabilities that he takes #1 or #100 are equal. Otherwise assume that he takes the  $n$ -th seat, where  $n$  is a number between 2 and 99. Everyone between 2 and  $(n-1)$  will get his own seat. That means the  $n$ -th passenger essentially becomes the new "drunk" guy with designated seat #1. If he chooses #1, all the rest of the passengers will have the correct seats. If he takes #100, then you will not get your seat. (The probabilities that he takes #1 or #100 are again equal.) Otherwise he will just make another passenger down the line the new "drunk" guy with designated seat #1 and each new "drunk" guy has equal probability of taking #1 or #100. Since at all jump points there's an equal probability for the "drunk" guy to choose seat #1 or 100, by symmetry, the probability that you, as the 100th passenger, will seat in #100 is 0.5.

## N points on a circle

Given  $N$  points drawn randomly on the circumference of a circle, what is the probability that they are all within a semicircle?<sup>4</sup>

*Solution:* Let's start at one point and clockwise label the points as  $1, 2, \dots, N$ . The probability that all the remaining  $N-1$  points from 2 to  $N$  are in the clockwise semicircle starting at point 1 (That is, if point 1 is at 12:00, points 2 to  $N$  are all between 12:00 and 6:00) is  $1/2^{N-1}$ . Similarly the probability that a clockwise semicircle starting at any point  $i$ , where  $i \in \{2, \dots, N\}$  contains all the other  $N-1$  points is also  $1/2^{N-1}$ .

Claim: the events that all the other  $N-1$  points are in the clockwise semicircle starting at point  $i$ ,  $i = 1, 2, \dots, N$  are mutually exclusive. In other words, if we, starting at point  $i$  and proceeding clockwise along the circle, sequentially encounter points  $i+1, i+2, \dots, N, 1, \dots, i-1$  in half a circle, then starting at any other point  $j$ , we cannot encounter all

---

<sup>4</sup> Hint: Consider the events that starting from a point  $n$ , you can reach all the rest of the points on the circle clockwise,  $n \in \{1, \dots, N\}$  in a semicircle. Are these events mutually exclusive?