1

Coin Tossing

1.1 Coin Tossing Probability Model

To start our journey into probability, we follow the exposition of Leo Breiman [9] by developing a coin tossing model. The experiment consists of tossing a fair coin n times and recording whether a head or tail comes up at each toss. Denote Ω_n to be the set of all possible outcomes, that is,

$$\Omega_n \triangleq \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) \mid \omega_i = H \text{ or } T \text{ for } i = 1, 2, \ldots, n\}. \tag{1.1}$$

The number of outcomes in Ω_n is 2^n as ω_1 can have either of the two values H or T, and for each of these values, ω_2 can have either of the two values H or T, and so on, to obtain

$$\text{Size of } \Omega_n = \underbrace{2 \cdot 2 \cdots 2}_{n \text{ times}} = 2^n.$$

Let $\omega \in \Omega_n$ be any particular outcome (n-long sequence of heads and tails), e.g., $\omega = (H, H, ..., H)$. Define

$$P_n(\omega) \triangleq \frac{1}{2^n}. \tag{1.2}$$

This is motivated by the intuition that any particular sequence of heads and tails is as likely as any other. We now make the following formal definition.

Definition 1 *Probability Space for Coin Tossing*

(1) Define the *sample space* Ω_n to be the set of all possible outcomes, i.e.,

$$\Omega_n \triangleq \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) \mid \omega_i = H \text{ or } T \text{ for } i = 1, 2, \ldots, n\}.$$

(2) Let $\mathcal{F}_n$ be the set of all subsets of Ω_n (which includes the empty set $\emptyset$ and Ω_n).

(3) Let $A \in \mathcal{F}_n$, i.e., any subset of Ω_n. Then the probability $P_n(A)$ is defined as the sum of the probability of all the outcomes in A, that is,

$$P_n(A) \triangleq \frac{\text{number of outcomes in A}}{2^n}. \tag{1.3}$$

Introduction to Probability Theory and Stochastic Processes by John Chiasson
©2013 John Wiley & Sons, Inc. Published 2013 by John Wiley & Sons, Inc.

COPYRIGHTED MATERIAL

Note that $P(\Omega_n) = 1$ and $P_n(\emptyset) = 0$ (there are no outcomes in the empty set).

Example 1 Consider $\Omega_2 \triangleq \{\omega = (\omega_1, \omega_2) \mid \omega_i = H \text{ or } T \text{ for } i = 1, 2\}$ which has $2^2 = 4$ outcomes, that is, $\Omega_2 = \{(H,H),(H,T),(T,H),(T,T)\}$. $\mathcal{F}_2$ is the class of all possible subsets of Ω_2, or explicitly,

$$\mathcal{F}_2 \triangleq \left\{ \begin{array}{c} \emptyset, \{(H,H)\}, \{(H,T)\}, \{(T,H)\}, \{(T,T)\}, \\ \{(H,H),(H,T)\}, \{(H,H),(T,H)\}, \{(H,H),(T,T)\}, \\ \{(H,T),(T,H)\}, \{(H,T),(T,T)\}, \{(T,H),(T,T)\}, \\ \{(H,H),(H,T),(T,H)\}, \{(H,H),(H,T),(T,T)\}, \\ \{(H,T),(T,H),(T,T)\}, \{(H,H),(T,H),(T,T)\}, \\ \{(H,H),(H,T),(T,H),(T,T)\} \end{array} \right\}.$$

For a single outcome $\omega \in \Omega_2$,

$$P_2(\omega) = \frac{1}{2^2} = \frac{1}{4}$$

so that, for example, with $\omega = (H,T)$,

$$P_2(\{(H,T)\}) = \frac{1}{4}.$$

More generally, for $A \in \mathcal{F}_2$, we define

$$P_2(A) \triangleq \frac{\text{number of } \omega \in A}{2^2}.$$

As another example, if $A = \{(H,H),(H,T),(T,H)\}$, then there are three outcomes in A and so $P_2(A) = 3/4$. The interpretation of this is that if we flip a coin twice, the probability that either $(H,H),(H,T)$, or (T,H) occurs is $3/4$.

Example 2 Let $A = \{(\omega_1, \omega_2, \ldots, \omega_n) \in \Omega_n \mid \omega_1 = H\}$. Then

$$P(A) = \frac{2^{n-1}}{2^n} = \frac{1}{2}$$

as there are 2^{n-1} outcomes in Ω_n that have $\omega_1 = H$. The interpretation is that a coin is flipped n times and the probability that heads appears on the first toss is $1/2$.

Similarly, $A^c = \{(\omega_1, \omega_2, \ldots, \omega_n) \in \Omega_n | \omega_1 = T\}$ has probability

$$P(A^c) = \frac{2^{n-1}}{2^n} = \frac{1}{2}$$

as there are also 2^{n-1} outcomes in Ω_n that have $\omega_1 = T$.

Properties of the Coin Tossing Probability Space

Observe that the coin tossing probability function P_n is defined on subsets of Ω_n, i.e., $P_n(A)$ is defined on all $A \in \mathcal{F}_n$ and takes values in the closed interval $[0,1]$. Further, P_n satisfied the following three properties:

I. $P_n(\Omega_n) = 1$.

II. $P_n(A) \geq 0$ for all $A \in \mathcal{F}_n$.

III. For $A, B \in \mathcal{F}_n$ with $A \cap B = \emptyset$, then

$$P_n(A \cup B) = P_n(A) + P_n(B).$$

The triple $(\Omega_n, \mathcal{F}_n, P_n)$ is called a *probability space* and $P_n \colon \mathcal{F}_n \to [0,1]$ is called a *probability distribution* function.

Remark We assume the reader is familiar with elementary set theory and set operations such as union, intersection, and complement. See the appendix at the end of this chapter on page 50 for some Venn diagrams illustrating set operations.

Remark The properties **I, II**, and **III** will later be used to define the axioms of a probability space. Many useful relations can be deduced from just these three properties. For example, by property **III** and property **I**, if $A \subset \Omega_n$, then $A \cup A^c = \Omega_n$ with $A \cap A^c = \emptyset$ so that $P_n(\Omega_n) = P_n(A \cup A^c) = P_n(A) + P_n(A^c)$ and therefore $P_n(A^c) = 1 - P_n(A)$. See Problem 10 for more examples.

1.2 Combinatorics: $\binom{n}{k}$

Let's now compute the probability of k heads in n tosses. This is just a problem in combinatorics. With $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ we are asking, How many ways can we choose k of the ω_i to be H and of course the $n-k$ remaining ω_i to be T? Well, there are n choices of the ω_i to be the first H; for each of these choices, there are $n-1$ remaining ω_i to be the second H; then $n-2$ remaining of the ω_i to be the third H,..., and finally the k^{th} head H can be chosen to be any of the $n-(k-1)$ remaining ω_i. Thus there are

$$n\,(n-1)\cdots(n-(k-1)) = \frac{n!}{(n-k)!} \tag{1.4}$$

ways of choosing k of the ω_i in $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ to be heads. However, any permutation of these chosen k positions ω_i for heads gives the same outcome, and there are $k!$ ways to permute the k positions we have chosen to be H. Hence there are

$$\binom{n}{k} \triangleq \frac{n!}{(n-k)!k!} \tag{1.5}$$

outcomes in Ω_n containing exactly k heads. As a result, the probability of k heads in n tosses is given by

$$P_n(\{\omega \in \Omega_n | \ \omega \text{ has } k \text{ heads}\}) = \binom{n}{k}\frac{1}{2^n} = \frac{n!}{(n-k)!k!}\frac{1}{2^n}. \tag{1.6}$$

We define $0! \triangleq 1$ so that

$$\binom{n}{0} = \frac{n!}{(n-0)!0!} = \frac{n!}{n!0!} = 1$$

and

$$\binom{n}{n} \triangleq \frac{n!}{(n-n)!n!} = \frac{n!}{0!n!} = 1.$$

Example 3 Suppose $n = 3$ and $k = 2$. Then by (1.4) there are $3 \cdot 2 = \dfrac{3!}{(3-2)!} = 6$ ways of choosing two of the ω_i to be heads and the remaining one to be tails. That is, one can choose ω_1, ω_2, or ω_3 to be the first heads (denote it as H_1), and for each these 3 choices there are 2 remaining ω_i as choices to put the second head (denote it as H_2). These six possibilities are

$$\begin{gathered}(H_1, H_2, T), (H_1, T, H_2),\\ (H_2, H_1, T), (T, H_1, H_2),\\ (H_2, T, H_1), (T, H_2, H_1).\end{gathered}$$

However, we do not want to consider (H_1, H_2, T) as a different outcome from (H_2, H_1, T), and so on. As there are $2! = 2$ ways to permute H_1 and H_2, we have

$$\binom{3}{2} \triangleq \frac{3!}{(3-2)!2!} = 3$$

ways of having two heads in three tosses. That is,

$$\{(H, H, T), (H, T, H), (T, H, H)\}.$$

Example 4 *The Binomial Formula*
The quantities $\binom{n}{k}$ for $k = 0, 1, ..., n$ form the coefficients in the binomial expansion of $(a+b)^n$, that is,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

This follows from the fact that the coefficient of $a^k b^{n-k}$ is just how many different ways one can choose "a" from k of the n factors in

$$\underbrace{(a+b)\cdots(a+b)}_{n} = (a+b)^n$$

and "b" from the remaining $n-k$ factors.

Example 5 *The Number of Subsets of a Set*

Another interpretation of $\binom{n}{k}$ is the number of ways of choosing k objects from n objects without regard to order (we do not care about the order the k objects are selected). Thus if $S \triangleq \{s_1, s_2, ..., s_N\}$ is a set with N objects, the number of ways we can pick k of these objects is $\binom{N}{k}$. Hence the number of subsets of S of size k is $\binom{N}{k}$. The total number of subsets of S is then

$$\binom{N}{0} + \binom{N}{1} + \binom{N}{2} + \cdots + \binom{N}{N} = \sum_{k=0}^{N} \binom{N}{k} 1^k 1^{N-k} = (1+1)^N = 2^N$$

as there is $\binom{N}{0} = 1$ subset with no elements (i.e., the empty set $\emptyset$), there are $\binom{N}{1} = N$ subsets with 1 element, etc.

In particular, consider the probability space $(\Omega_2, \mathcal{F}_2, P_2)$ in Example 1 where $\Omega_2 = \{(H,H),(H,T),(T,H),(T,T)\}$ and $\mathcal{F}_2$ consists of all possible subsets of Ω_2. Then, as Ω_2 has four elements, the number of subsets of Ω_2 is $2^4 = 16$ and these sixteen subsets making up $\mathcal{F}_2$ are explicitly listed in Example 1.

Example 6 *Stirling's Approximation* [9][14]

Suppose we flip a coin $2n$ times. What is the probability of n heads in the $2n$ tosses? We have

$$P_{2n}(\{\omega \in \Omega_{2n} | \ n \text{ heads in } 2n \text{ tosses}\}) = \binom{2n}{n} 2^{-2n} = \frac{(2n)!}{n!n!} 2^{-2n}. \quad (1.7)$$

In Problem 2, the reader is asked to show that $\binom{2n}{k} 2^{-2n}$ is a maximum for $k = n$, making it by definition the *most likely value* for the number of heads in $2n$ tosses. Stirling's approximation (see [9] or [14]) states that

$$n! = e^{-n} n^n \sqrt{2\pi n}(1 + \epsilon_n),$$

where $\epsilon_n \to 0$ as $n \to \infty$, that is,

$$\frac{n!}{e^{-n}n^n\sqrt{2\pi n}} \overset{n\to\infty}{\to} 1. \tag{1.8}$$

Using Stirling's approximation, we have

$$\begin{aligned}
\frac{(2n)!}{n!n!}2^{-2n} &= \frac{e^{-2n}(2n)^{2n}\sqrt{2\pi 2n}(1+\epsilon_{2n})}{e^{-n}n^n\sqrt{2\pi n}(1+\epsilon_n)e^{-n}n^n\sqrt{2\pi n}(1+\epsilon_n)}2^{-2n} \\
&= \frac{e^{-2n}(2n)^{2n}\sqrt{2\pi 2n}}{e^{-2n}n^{2n}2\pi n}2^{-2n}\frac{1+\epsilon_{2n}}{(1+\epsilon_n)^2} \\
&= \frac{1}{\sqrt{\pi n}}\frac{1+\epsilon_{2n}}{(1+\epsilon_n)^2} \\
&\to \frac{1}{\sqrt{\pi n}} \to 0 \text{ as } n \to \infty.
\end{aligned} \tag{1.9}$$

The conclusion of this analysis is that the proportion of sequences of $2n$ tosses of a coin such that heads comes up *exactly* n times goes to zero. See Figure 1.1. Loosely speaking, in 1000 tosses of a fair coin, it is *not* reasonable to expect *exactly* 500 heads. (Note that $1/\sqrt{\pi\,(500)} \approx 0.025$.)

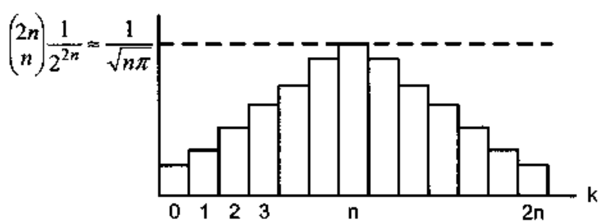

FIGURE 1.1. $\binom{2n}{k}\frac{1}{2^{2n}}$ versus k.

1.2.1 Multinomial Formula

The binomial formula has been discussed, and more generally there is the multinomial formula. This arises when one considers the problem of how many different ways a set S of size n can be partitioned into r subsets of size $k_1, k_2, ..., k_r$, respectively, where the k_i are positive integers with $k_1 + k_2 + \cdots + k_r = n$. Let's clarify what this means by an example.

Example 7 *Partitioning a Set* [15]
Let $S = \{1, 2, 3, 4\}$ be a set containing four elements. How many ways can we partition S of size $n = 4$ into 3 subsets of size $k_1 = 1, k_2 = 2, k_3 = 1$?

The 12 possible partitions are

$$\begin{array}{ll}
\{1\},\{2,3\},\{4\} & \{2\},\{1,3\},\{4\} \\
\{1\},\{2,4\},\{3\} & \{2\},\{1,4\},\{3\} \\
\{1\},\{3,2\},\{2\} & \{2\},\{3,4\},\{1\} \\
\\
\{3\},\{1,2\},\{4\} & \{4\},\{1,2\},\{3\} \\
\{3\},\{1,4\},\{2\} & \{4\},\{1,3\},\{2\} \\
\{3\},\{2,4\},\{1\} & \{4\},\{2,3\},\{1\}.
\end{array}$$

Note that within a subset of a particular partition the order of the elements does not matter. For example, the partition $\{1\},\{2,3\},\{4\}$ is considered the same as $\{1\},\{3,2\},\{4\}$.

Going back to the general case, we want to partition a set S of size n into r subsets of size $k_1, k_2, ..., k_r$ with $k_1+k_2+\cdots+k_r = n$. Then there are $\binom{n}{k_1}$ ways of choosing k_1 elements for the first subset of the partition. For each of these $\binom{n}{k_1}$ choices, there are $\binom{n-k_1}{k_2}$ ways of choosing k_2 elements for the second subset of the partition. Thus there are a total of $\binom{n}{k_1}\binom{n-k_1}{k_2}$ ways of choosing k_1 elements for the first subset and k_2 elements for the second subset. Continuing in this manner for all r subsets, we have

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{\overbrace{n-(k_1+k_2+\cdots+k_{r-1})}^{k_r}}{k_r}$$

possible ways of partitioning a set S of size n into r subsets where the first subset has size k_1, the second subset has size k_2, and so on, with $k_1+k_2+\cdots+k_r = n$. Using the definition of $\binom{n}{k}$, it follows immediately that

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

For convenience of notation, we define the *multinomial coefficients*

$$\binom{n}{k_1\, k_2\, \cdots\, k_r} \triangleq \frac{n!}{k_1!k_2!\cdots k_r!}. \tag{1.10}$$

Example 8 *Partitioning a Set (continued)* [15]

With $S = \{1,2,3,4\}$ so $n = 4$ and choosing $k_1 = 1, k_2 = 2, k_3 = 1$ we have

$$\binom{4}{1\,2\,1} \triangleq \frac{4!}{1!2!1!} = \frac{24}{2} = 12.$$

These are the 12 partitions explicitly given above in Example 7.

The multinomial (trinomial) formula is

$$(a_1 + a_2 + a_3)^n = \underbrace{\sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \sum_{k_3=1}^{n}}_{k_1+k_2+k_3=n} \binom{n}{k_1\, k_2\, k_3} a_1^{k_1} a_1^{k_2} a_3^{k_3}.$$

This follows because the coefficient of $a_1^{k_1} a_1^{k_2} a_3^{k_3}$ in the expansion of

$$(a_1 + a_2 + a_3)^n = \underbrace{(a_1 + a_2 + a_3) \cdots (a_1 + a_2 + a_3)}_{n \text{ times}}$$

is the number of ways one can choose a_1 from k_1 of the n factors, a_2 from k_2 of the remaining $n - k_1$ factors, and finally a_3 from the last remaining $k_3 = n - (k_1 + k_2)$ factors, which is just $\binom{n}{k_1\, k_2\, k_3}$. In general we have

$$(a_1 + a_2 + \cdots + a_r)^n = \underbrace{\sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_r=1}^{n}}_{k_1+k_2+\cdots+k_r=n} \binom{n}{k_1\, k_2\, \cdots\, k_r} a_1^{k_1} a_2^{k_2} \cdots a_r^{k_r}.$$

*1.2.2 Occupancy Problems**[1]

Suppose we have r identical balls and we want to put them into n bins. How many different ways can we do this? Let r_1 denotes the number of balls in bin 1, r_2 denotes the number of balls in bin 2, etc., where $r_i \geq 0$ for $i = 1, ..., n$. Note that r_i can be zero corresponding to putting none of the balls in the i^{th} bin. We are therefore asking how many different ways can we choose n nonnegative integers such that

$$r_1 + r_2 + \cdots + r_n = r?$$

A particularly nice way of showing how to compute the solution to this question is given by W. Feller (see page 38 of [14]). Consider, for example, we have 12 balls and we want to put them in 6 bins as shown in the top half of Figure 1.2. The bottom half of Figure 1.2 is an equivalent representation of the $n = 6$ bins and $r = 12$ balls.

[1]Sections marked with an asterisk may be skipped without loss of continuity.

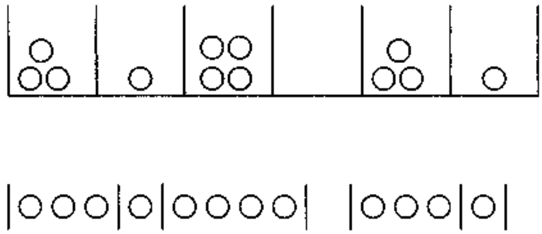

FIGURE 1.2. $r = 12$ balls distributed among $n = 6$ bins.

Note that in the bottom half of Figure 1.2 there are $n + 1 = 7$ bars to delineate the $n = 6$ bins. The key observation here is that the two end bars in this representation must be fixed; but within the two end bars, the $r = 12$ balls and remaining $n - 1 = 5$ bars can be in *any* order. That is, there are $n-1+r$ spots (a spot is either a ball or a bar) to put the $n-1=5$ bars and the $r = 12$ balls. Thus we are asking how many different ways can we put $n - 1$ bars in $n - 1 + r$ spots (with of course the balls going in the remaining spots). This is simply

$$\binom{n-1+r}{n-1}.$$

This is also the same as asking how many different ways can we put r balls in $n - 1 + r$ spots (with of course the bars going in the remaining spots) so that

$$\binom{n-1+r}{r} = \binom{n-1+r}{n-1}.$$

Of course one can straightforwardly show these two expressions are equal by direct calculation.

Example 9 *Throwing r Dice*

Suppose one throws r dice. Then how many different ways are there for the faces of the dice to land? Let r_1 denotes the number of dice that land face up with a 1 (bin 1), r_2 denotes the number of dice that land face up with a 2 (bin 2), etc. The problem is equivalent to how many different ways can we choose six nonnegative integers $r_i \geq 0$ with

$$r_1 + r_2 + \cdots + r_6 = r?$$

As there are $n = 6$ faces (bins) and r dice (balls), the answer is

$$\binom{n-1+r}{n-1} = \binom{5+r}{5}.$$

Example 10 *The Number of Terms in a Multinomial Expansion*
By the multinomial theorem, we have the expansion

$$(a_1 + a_2 + \cdots + a_n)^r = \sum_{\substack{r_i \geq 0 \\ r_1+r_2+\cdots+r_n=r}} \binom{r}{r_1\, r_2\, \cdots\, r_n} a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n},$$

where the multinomial coefficients are

$$\binom{r}{r_1\, r_2\, \cdots\, r_n} \triangleq \frac{r!}{r_1! r_2! \cdots r_r!}.$$

As the sum is over all possible values of $r_i \geq 0$ satisfying $r_1+r_2+\cdots+r_n = r$, there are thus

$$\binom{n-1+r}{r}$$

many terms (multinomial coefficients) in the expansion.

1.3 Random Variables on Ω_n

Functions from the set of outcomes to the real numbers will be invaluable in studying probability. For our current model of coin tossing, we have the following definition.

Definition 2 Any function from Ω_n to the real numbers $\mathbb{R}$ is called a *random variable.*

Example 11 *Coin Tossing Random Variables*
Let $X_i(\omega) : \Omega_n \to \mathbb{R}$

$$X_i(\omega) = \begin{cases} 0, & \omega_i = T, \\ 1, & \omega_i = H. \end{cases} \tag{1.11}$$

In words, X_i is 0 if the i^{th} toss is tails and is 1 is the i^{th} toss is heads.

The set $A_i \triangleq \{\omega \in \Omega_n | \ X_i(\omega) = 1\}$ is the set of all $\omega \in \Omega_n$ such that the i^{th} toss is heads, and its probability is

$$P_n(A_i) = P_n(\{\omega \in \Omega_n | \ X_i(\omega) = 1\}) = 2^{n-1}\frac{1}{2^n} = \frac{1}{2} \tag{1.12}$$

as there are 2^{n-1} outcomes with the i^{th} toss being heads. The set

$$A_1 \cap A_2 \cap \cdots \cap A_n = \{\omega \in \Omega_n | X_1(\omega) = 1, \ldots, X_n(\omega) = 1\}$$

is the set of all $\omega \in \Omega_n$ such that *all* tosses are heads. Its probability is

$$P_n(A_1 \cap A_2 \cap \cdots \cap A_n) = P_n(\{(H, H, ..., H)\}) = \frac{1}{2^n}$$

as there is only one outcome in Ω_n with all heads. We have thus shown by direct computation that

$$P_n(A_1 \cap A_2 \cap \cdots \cap A_n) = P_n(A_1)P_n(A_2) \cdots P_n(A_n).$$

Example 12 *Coin Tossing Random Variables*
Continuing with Example 11, we show ($n \geq 3$)

$$P_n(A_1^c \cap A_3) = P_n(A_1^c)P_n(A_3).$$

We have

$$\begin{aligned}
A_1^c &= \{\omega \in \Omega_n | X_1(\omega) = 1\}^c = \{\omega \in \Omega_n | X_1(\omega) = 0\}, \\
A_3 &\triangleq \{\omega \in \Omega_n | X_3(\omega) = 1\}, \\
A_1^c \cap A_3 &= \{\omega \in \Omega_n | X_1(\omega) = 0, X_3(\omega) = 1\}.
\end{aligned}$$

There are 2^{n-1} outcomes in A_1^c so that $P_n(A_1^c) = 2^{n-1}/2^n = 1/2$. The set A_3 has 2^{n-1} outcomes so that $P_n(A_3) = 2^{n-1}/2^n = 1/2$. Finally, the set $A_1^c \cap A_3$ has 2^{n-2} outcomes so that its probability is given by $P_n(A_1^c \cap A_3) = 2^{n-2}/2^n = 1/4$. Thus we have

$$P_n(A_1^c \cap A_3) = P_n(A_1^c)P_n(A_3).$$

Similarly, the reader can show that for any combination of *distinct* $A_{i_1}, ..., A_{i_k}$ ($i_1 \neq i_2 \neq \cdots \neq i_n$)

$$P_n(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P_n(A_{i_1})P_n(A_{i_2}) \cdots P_n(A_{i_k}).$$

In this case we say the A_i are *independent* sets.

Example 13 *Independent Sets*
We say the sets A_1, A_2, A_3 are independent if and only if

$$\begin{aligned}
P_n(A_1 \cap A_2) &= P_n(A_1)P_n(A_2), \\
P_n(A_1 \cap A_3) &= P_n(A_1)P_n(A_3), \\
P_n(A_2 \cap A_3) &= P_n(A_2)P_n(A_3), \\
P_n(A_1 \cap A_2 \cap A_3) &= P_n(A_1)P_n(A_2)P_n(A_3).
\end{aligned}$$

As explained in the previous example, it follows that the sets A_1, A_2, A_3 are independent.

Notation 1 *Indices*

Typically, the letters $i, j, k, ...$ are used for indices. For example, the three sets A_1, A_2, A_3 are independent if $P_n(A_i \cap A_j) = P_n(A_i)P_n(A_j)$ for $i, j = 1, 2, 3$ and $i \neq j$ *and* $P_n(A_1 \cap A_2 \cap A_3) = P_n(A_1)P_n(A_2)P_n(A_3)$. The problem with this notation is that the we soon run out of letters. For suppose we have n sets $A_1, A_2, ..., A_n$ where n is some huge number. If we select two of the sets, then we can denote them as A_i and A_j where $1 \leq i, j \leq n$. However, if we select k sets (k not specified), then how do we choose k letters $i, j, ...$ if $k > 26$? To get around this, we just use a different notation by letting $i_1, i_2, ..., i_k$ be the k indices. Then we say that A_1, A_2, A_3 are independent if $P_n(A_{i_1} \cap A_{i_2}) = P_n(A_{i_1})P_n(A_{i_2})$ for $i_1, i_2 = 1, 2, 3$ with $i_1 \neq i_2$ *and* $P_n(A_1 \cap A_2 \cap A_3) = P_n(A_1)P_n(A_2)P_n(A_3)$ (here $i_1 \leftrightarrow i$ and $i_2 \leftrightarrow j$). We treat k sets by simply using the indices $i_1, i_2, ..., i_k$ with $1 \leq i_1 \leq n, 1 \leq i_2 \leq n, ..., 1 \leq i_k \leq n$.

Notation 2 *Set Notation*

Up to now we have been using the notation $\{\omega \in \Omega_n | S_n(\omega) = k\}$ to describe a set of outcomes in Ω_n specified by the random variable S_n. There are various ways to shorten up this notation. Specifically,

$$\{\omega \in \Omega_n | S_n(\omega) = k\}, \{\omega | S_n(\omega) = k\}, \{S_n(\omega) = k\}, \text{ or } \{S_n = k\}$$

all mean the same thing. In this chapter we have used only $\{\omega \in \Omega_n | S_n(\omega) = k\}$. In later chapters we will tend to use $\{\omega | S_n(\omega) = k\}$. In terms of the probability of a set, one writes

$$\begin{aligned} P_n(\{\omega \in \Omega_n | S_n(\omega) = k\}) &= P_n(\{\omega | S_n(\omega) = k\}) \\ &= P_n(\{S_n(\omega) = k\}) \\ &= P_n(\{S_n = k\}) \\ &= P_n(S_n = k). \end{aligned}$$

The last expression $P_n(S_n = k)$ tends to be the notation in most books as it is very short to write down.

1.4 Binomial Probability Distribution

Let's continue with our coin tossing probability space $(\Omega_n, \mathcal{F}_n, P_n)$ and define a new random variable on Ω_n by

$$S_n(\omega) \triangleq X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega), \tag{1.13}$$

where as before, for $i = 1, ..., n$ we define

$$X_i(\omega) = \begin{cases} 0, & \omega_i = T, \\ 1, & \omega_i = H. \end{cases} \tag{1.14}$$

In words, $S_n(\omega)$ is the number of heads in n tosses, i.e., the number of heads in the outcome $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$.

The set

$$\{\omega \in \Omega_n |\ S_n(\omega) = k\}$$

is the set of all outcomes $(\omega_1, \omega_2, \ldots, \omega_n)$ such that there are k heads in n tosses. By (1.6) we know

$$\begin{aligned} P_n(\{\omega \in \Omega_n |\ S_n(\omega) = k\}) &= \frac{\text{number of } \omega \in \Omega_n \text{ with } S_n(\omega) = k}{2^n} \\ &= \binom{n}{k} \frac{1}{2^n}. \end{aligned}$$

For all $\omega \in \Omega_n$, $S_n(\omega)$ only takes on values in the set $\Omega_{S_n} \triangleq \{0, 1, 2, ..., n\}$, which is the *range* of the random variable (function) S_n. For $k \in \Omega_{S_n}$, define

$$P_{S_n}(k) \triangleq P_n(\{\omega \in \Omega_n | S_n(\omega) = k\}) = \binom{n}{k} \frac{1}{2^n}, \tag{1.15}$$

where

$$\begin{aligned} P_{S_n}(k) &\geq 0, \\ \sum_{k=0}^{n} P_{S_n}(k) &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{2^n} = 1. \end{aligned}$$

Let $\mathcal{F}_{S_n}$ denote all possible subsets of Ω_{S_n}, and define for any $A \in \mathcal{F}_{S_n}$

$$P_{S_n}(A) \triangleq \sum_{k \in A} P_{S_n}(k) \geq 0.$$

The reader can show that P_{S_n} satisfies the following three properties:

I. $P_{S_n}(\Omega_{S_n}) = 1$.

II. $P_{S_n}(A) \geq 0$ for all $A \in \mathcal{F}_{S_n}$.

III. For all $A, B \in \mathcal{F}_{S_n}$ with $A \cap B = \emptyset$

$$P_{S_n}(A \cup B) = P_{S_n}(A) + P_{S_n}(B).$$

We then say that $(\Omega_{S_n}, \mathcal{F}_{S_n}, P_{S_n})$ is a *probability space*. The probability function $P_{S_n}(k)$ in (1.15) is called the *binomial* probability distribution function and is denoted as $\mathcal{B}(n, 1/2)$ where $\mathcal{B}$ stands for binomial, n is the number of coin tosses, and $1/2$ is the probability of heads. We also say that the random variable S_n defined on the probability space $(\Omega_n, \mathcal{F}_n, P_n)$ *induces* the binomial probability distribution on its range. More precisely, it induces the probability space $(\Omega_{S_n}, \mathcal{F}_{S_n}, P_{S_n})$.

Example 14 *Binomial Probability Distribution* $(\Omega_{S_2}, \mathcal{F}_{S_2}, P_{S_2})$
Consider the probability space $(\Omega_{S_2}, \mathcal{F}_{S_2}, P_{S_2})$ where $\Omega_{S_2} \triangleq \{0, 1, 2\}$. Then let

$$\begin{aligned} \mathcal{F}_{S_2} &\triangleq \{\text{all possible subsets of } \Omega_{S_2}\} \\ &= \Big\{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\Big\} \end{aligned}$$

and define

$$\begin{aligned} P_{S_2}(0) &= \binom{2}{0}\frac{1}{2^2} = \frac{1}{4}, \\ P_{S_2}(1) &= \binom{2}{1}\frac{1}{2^2} = \frac{1}{2}, \\ P_{S_2}(2) &= \binom{2}{2}\frac{1}{2^2} = \frac{1}{4}. \end{aligned}$$

Further, for any $A \in \mathcal{F}_{S_2}$, define $P_{S_2}(A)$ as

$$P_{S_2}(A) \triangleq \sum_{k \in A} P_{S_2}(k).$$

For example, let $A = \{0, 2\} \in \mathcal{F}_{S_2}$, then

$$P_{S_2}(A) = P_{S_2}(\{0,2\}) = P_{S_2}(\{0\}) + P_{S_2}(\{2\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Definition 3 *Expected Value*
Let $X(\omega)$ be a random variable on Ω_n. Define the *expected value* of X as

$$E[X] \triangleq \sum_{\omega \in \Omega_n} X(\omega) P_n(\omega) = \sum_{\omega \in \Omega_n} X(\omega)\frac{1}{2^n}.$$

Example 15 *Expected Value of the Coin Tossing RVs*
Consider the coin tossing random variables on $(\Omega_n, \mathcal{F}_n, P_n)$, that is,

$$X_i(\omega) = \begin{cases} 0, & \omega_i = T, \\ 1, & \omega_i = H. \end{cases}$$

Then

$$\begin{aligned}
E[X_i] &= \sum_{\omega\in\Omega_n} X_i(\omega)P_n(\omega) \\
&= \sum_{\{\omega\in\Omega_n|\ X_i(\omega)=0\}} X_i(\omega)P_n(\omega) + \sum_{\{\omega\in\Omega_n|\ X_i(\omega)=1\}} X_i(\omega)P_n(\omega) \\
&= \sum_{\{\omega\in\Omega_n|\ X_i(\omega)=0\}} 0\cdot P_n(\omega) + \sum_{\{\omega\in\Omega_n|\ X_i(\omega)=1\}} 1\cdot P_n(\omega) \\
&= 0\cdot\Big(\sum_{\{\omega\in\Omega_n|\ X_i(\omega)=0\}} P_n(\omega)\Big) + 1\cdot\Big(\sum_{\{\omega\in\Omega_n|\ X_i(\omega)=1\}} P_n(\omega)\Big) \\
&= 0\cdot P_n(\{\omega\in\Omega_n|\ X_i(\omega)=0\}) + 1\cdot P_n(\{\omega\in\Omega_n|\ X_i(\omega)=1\}) \\
&= \frac{1}{2},
\end{aligned}$$

where $P_n(\{\omega\in\Omega_n|\ X_i(\omega)=0\}) = P_n(\{\omega\in\Omega_n|\ X_i(\omega)=1\}) = 1/2$ (see Example 2).

We now use this result to compute $E[S_n]$ as follows:

$$\begin{aligned}
E[S_n] &\triangleq \sum_{\omega\in\Omega_n} S_n(\omega)P_n(\omega) \\
&= \sum_{\omega\in\Omega_n} (X_1(\omega)+X_2(\omega)+\cdots+X_n(\omega))\frac{1}{2^n} \\
&= \sum_{\omega\in\Omega_n} X_1(\omega)\frac{1}{2^n} + \sum_{\omega\in\Omega_n} X_2(\omega)\frac{1}{2^n} + \cdots + \sum_{\omega\in\Omega_n} X_n(\omega)\frac{1}{2^n} \\
&= \sum_{i=1}^{n} E[X_i] \\
&= n\frac{1}{2}.
\end{aligned}$$

Example 16 *Expected Value of S_n*

We now recompute $E[S_n]$ using a longer but still instructive procedure. First we decompose Ω_n into a union of disjoint sets given by

$$\Omega_n = \{\omega\in\Omega_n|S_n(\omega)=0\}\cup\{\omega\in\Omega_n|S_n(\omega)=1\}\cup\cdots\cup\{\omega\in\Omega_n|S_n(\omega)=n\}.$$

As shown above,

$$\begin{aligned}
P_n(\{\omega\in\Omega_n|S_n(\omega)=k\}) &= \frac{\text{number of } \omega\in\Omega_n \text{ with } S_n(\omega)=k}{2^n} \\
&= \binom{n}{k}\frac{1}{2^n}.
\end{aligned}$$

Then $E[S_n]$ is computed as follows:

$$
\begin{aligned}
E[S_n] &\triangleq \sum_{\omega\in\Omega_n} S_n(\omega)P_n(\omega) \\
&= \sum_{\{\omega|S_n(\omega)=0\}} S_n(\omega)P_n(\omega) + \sum_{\{\omega|S_n(\omega)=1\}} S_n(\omega)P_n(\omega) + \cdots \\
&\quad + \sum_{\{\omega|S_n(\omega)=n\}} S_n(\omega)P_n(\omega) \\
&= \sum_{\{\omega|S_n(\omega)=0\}} 0\cdot P_n(\omega) + \sum_{\{\omega|S_n(\omega)=1\}} 1\cdot P_n(\omega) + \cdots \\
&\quad + \sum_{\{\omega|S_n(\omega)=n\}} n\cdot P_n(\omega) \\
&= 0\cdot\Big(\sum_{\{\omega|S_n(\omega)=0\}} P_n(\omega)\Big) + 1\cdot\Big(\sum_{\{\omega|S_n(\omega)=1\}} P_n(\omega)\Big) + \cdots \\
&\quad + n\cdot\Big(\sum_{\{\omega|S_n(\omega)=n\}} P_n(\omega)\Big) \\
&= 0\cdot P_n(\{\omega\in\Omega_n|S_n(\omega)=0\}) + 1\cdot P_n(\{\omega\in\Omega_n|S_n(\omega)=1\}) + \cdots \\
&\quad + n\cdot P_n(\{\omega\in\Omega_n|S_n(\omega)=n\}) \\
&= 0\cdot\binom{n}{0}\frac{1}{2^n} + 1\cdot\binom{n}{1}\frac{1}{2^n} + \cdots + n\cdot\binom{n}{n}\frac{1}{2^n} \\
&= \sum_{k=0}^{n} k\binom{n}{k}\frac{1}{2^n} \\
&= \frac{n}{2},
\end{aligned}
$$

where $\sum_{k=0}^{n} k\binom{n}{k}\frac{1}{2^n} = \frac{n}{2}$ follows by Problem 1.

1.5 Weak Law of Large Numbers for Coin Tossing

Before stating and proving the weak law of large numbers, we first prove Chebyshev's inequality.

Theorem 1 *Chebyshev's Inequality*
For any random variable Z defined on Ω_n and any $\epsilon > 0$, we have

$$P_n(\{\omega \in \Omega_n |\ |Z(\omega)| \geq \epsilon\}) \leq \frac{1}{\epsilon^2} E[Z^2].$$

Proof.

$$\begin{aligned}
P_n(\{\omega \in \Omega_n |\ |Z(\omega)| \geq \epsilon\}) &= (\text{no. } \omega \in \Omega_n \text{ such that } |Z(\omega)| \geq \epsilon) \frac{1}{2^n} \\
&= \sum_{\{\omega \in \Omega_n |\, |Z(\omega)| \geq \epsilon\}} 1 \cdot \frac{1}{2^n} \\
&\leq \sum_{\{\omega \in \Omega_n |\, |Z(\omega)| \geq \epsilon\}} \frac{Z^2(\omega)}{\epsilon^2} \frac{1}{2^n} \\
&\leq \sum_{\{\omega \in \Omega_n\}} \frac{Z^2(\omega)}{\epsilon^2} \frac{1}{2^n} \\
&= \frac{1}{\epsilon^2} \sum_{\{\omega \in \Omega_n\}} Z^2(\omega) \frac{1}{2^n} \\
&= \frac{1}{\epsilon^2} E[Z^2].
\end{aligned}$$

■

We now use Chebyshev's inequality to prove the *Weak Law of Large Numbers (WLLN)*.

Theorem 2 *Weak Law of Large Numbers (WLLN)*
The weak law of large numbers states that given any $\epsilon > 0$ (no matter how small), we have

$$\lim_{n \to \infty} P_n(\{\omega \in \Omega_n |\ |S_n(\omega)/n - 1/2| < \epsilon\}) = 1. \tag{1.16}$$

Proof. By Chebyshev's Inequality and Problem 4, we obtain

$$P_n(\{\omega \in \Omega_n |\ |S_n(\omega)/n - 1/2| \geq \epsilon\}) \leq \frac{1}{\epsilon^2} E[(S_n(\omega)/n - 1/2)^2] = \frac{1}{\epsilon^2} \frac{1}{4n}.$$

Ω_n is now written as the union of two *disjoint* sets as

$$\Omega_n = \{\omega \in \Omega_n |\ |S_n(\omega)/n - 1/2| < \epsilon\} \cup \{\omega \in \Omega_n |\ |S_n(\omega)/n - 1/2| \geq \epsilon\},$$

from which it follows that

$$\begin{aligned}1 &= P_n(\Omega_n)\\ &= P_n(\{\omega\in\Omega_n|\ |S_n(\omega)/n-1/2|<\epsilon\})+P_n(\{\omega\in\Omega_n|\ |S_n(\omega)/n-1/2|\geq\epsilon\}).\end{aligned}$$

Rearranging, we have

$$\begin{aligned}P_n(\{\omega\in\Omega_n|\ |S_n(\omega)/n-1/2|<\epsilon\}) &= 1-P_n(\{\omega\in\Omega_n|\ |S_n(\omega)/n-1/2|\geq\epsilon\})\\ &\geq 1-\frac{1}{\epsilon^2}\frac{1}{4n}.\end{aligned}$$

Letting $n\to\infty$ gives

$$\lim_{n\to\infty}P_n(\{\omega\in\Omega_n|\ |S_n(\omega)/n-1/2|<\epsilon\})=1.$$

■

Example 17 *Weak Law of Large Numbers*
Let $n=10^6$ and $\epsilon=1/100$ so that

$$\begin{aligned}P_{10^6}(\{\omega\in\Omega_{10^6}|\ |S_{10^6}(\omega)/10^6-1/2|<1/100\}) &\geq 1-\frac{1}{\epsilon^2}\frac{1}{4n}\bigg|_{n=10^6,\ \epsilon=10^{-2}}\\ &= 1-\frac{1}{400}\\ &= \frac{399}{400}.\end{aligned}$$

So, if you toss the coin one million times and count the number of heads, then the probability is at least 399/400 that the *ratio* $S_{10^6}(\omega)/10^6$ of the number of heads to the total number of tosses equals 1/2 within ± 0.01.

1.6 Consistency Between Probability Spaces

For any $n>0$, we have developed the probability spaces $(\Omega_n,\mathcal{F}_n,P_n)$ to model n tosses of a fair coin, that is, tossing a coin n times with the probability of heads on any given toss being 1/2. Specifically,

$$\Omega_n=\{(\omega_1,\omega_2,...,\omega_n)\mid \omega_i=H \text{ or } T \text{ for } i=1,2,...,n\}$$

and for each $\omega\in\Omega_n$

$$P_n(\omega)\triangleq\frac{1}{2^n}.$$

For $i=1,2,...,n$ we can define on *each* sample space Ω_n the random variables

$$X_i(\omega)=\begin{cases}1, & \omega_i=H,\\ 0, & \omega_i=T.\end{cases}$$

However, the random variables X_1, X_2 defined on Ω_2 are *different* than the random variables X_1, X_2 defined on Ω_3 as the former are defined on 2-tuples while the latter are defined on 3-tuples.

Similarly, for $k = 0, 1, ..., n$ we define on each sample space Ω_n the random variables

$$S_k(\omega) = X_1(\omega) + \cdots + X_k(\omega).$$

With $k < n$, let's now compare the two probability spaces $(\Omega_k, \mathcal{F}_k, P_k)$ and $(\Omega_n, \mathcal{F}_n, P_n)$. Specifically, consider the sets

$$\{\omega \in \Omega_k |\ S_k(\omega) = \ell\}$$

and

$$\{\omega \in \Omega_n |\ S_k(\omega) = \ell\},$$

where $\ell \leq k < n$. These sets are *not* the same as the outcomes in $\{\omega \in \Omega_k |\ S_k(\omega) = \ell\}$ are k-long while the outcomes in $\{\omega \in \Omega_n |\ S_k(\omega) = \ell\}$ are n-long. However the probabilities of these sets are equal, that is,

$$P_k(\{\omega \in \Omega_k |\ S_k(\omega) = \ell\}) = P_n(\{\omega \in \Omega_n |\ S_k(\omega) = \ell\}).$$

To show this, recall that we have already shown

$$P_k(\{\omega \in \Omega_k |\ S_k(\omega) = \ell\}) = \binom{k}{\ell} 2^{-k}.$$

To compute $P_n(\{\omega \in \Omega_n |\ S_k(\omega) = \ell\})$, note that there are $\binom{k}{\ell}$ ways to have ℓ heads on the first toss k tosses and tails on the next $k - \ell$ tosses. And, for each of these, there are 2^{n-k} ways of having either a head or a tail on the remaining $n - k$ tosses. Thus there are

$$\binom{k}{\ell} 2^{n-k}$$

outcomes in $\{\omega \in \Omega_n |\ S_k(\omega) = \ell\}$. Each of these outcomes have probability $1/2^n$ resulting in

$$\begin{aligned} P_n(\{\omega \in \Omega_n |\ S_k(\omega) = \ell\}) &= \binom{k}{\ell} 2^{n-k} \frac{1}{2^n} \\ &= \binom{k}{\ell} 2^{-k} \\ &= P_k(\{\omega \in \Omega_k |\ S_k(\omega) = \ell\}). \end{aligned}$$

To repeat, the two sets $\{\omega \in \Omega_k |\ S_k(\omega) = \ell\}$ and $\{\omega \in \Omega_n |\ S_k(\omega) = \ell\}$ are events from *different* probability spaces, but they model the same physical event which is getting ℓ heads in the first k tosses of a fair coin. As just shown, the two probability spaces $(\Omega_k, \mathcal{F}_k, P_k)$ and $(\Omega_n, \mathcal{F}_n, P_n)$ give the same value for the probability of this physical event. We say they are *consistent* probability spaces because they give the same probabilities when modeling the same physical event.

1.7 Random Walks: Reflection Principle*

We now use the simple coin tossing probability model to study random walks. This material is from Chapter 3 of Volume 1 of the classic work of W. Feller [14]. As above, we are considering the probability space $(\Omega_n, \mathcal{F}_n, P_n)$. For $i = 1, 2, ..., n$, define the random variables (functions) Y_i by

$$Y_i(\omega) = \begin{cases} +1, & \omega_i = H, \\ -1, & \omega_i = T \end{cases}$$

and define the random variable D_k by

$$D_k(\omega) \triangleq \sum_{i=1}^{k} Y_i(\omega).$$

Example 18 $D_k(\omega)$
For $\omega = (H, T, T, T) \in \Omega_4$, we have

$$\begin{aligned} Y_1(\omega) &= +1, \\ Y_2(\omega) &= -1, \\ Y_3(\omega) &= -1, \\ Y_4(\omega) &= -1, \end{aligned}$$

so that

$$\begin{aligned} D_1(\omega) &= 1, \\ D_2(\omega) &= 1 + (-1) = 0, \\ D_3(\omega) &= 1 + (-1) + (-1) = -1, \\ D_4(\omega) &= 1 + (-1) + (-1) + (-1) = -2. \end{aligned}$$

Definition 4 *Path*
Let $n > 0$ and r be integers. A *path* from the origin $(0, 0)$ to the point (n, r) is a polygonal line whose vertices have abscissas $1, ..., n$ and ordinates $(d_1, d_2, ..., d_n)$ with $d_n = r$. For convenience, we will refer to $(d_1, d_2, ..., d_n)$ as the path.

Example 19 *Path*
Consider the path $(d_1, d_2, d_3, d_4) = (1, 0, -1, -2)$ as in the previous example where $(n, r) = (4, -2)$. The path is drawn in Figure 1.3.

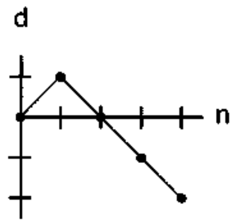

FIGURE 1.3. The path for the outcome $\omega = (H, T, T, T)$.

Let

$$\begin{aligned} n_H \text{ be the number of heads in } \omega &= (\omega_1, \omega_2, ..., \omega_n), \\ n_T \text{ be the number of tails in } \omega &= (\omega_1, \omega_2, ..., \omega_n). \end{aligned}$$

Then

$$\begin{aligned} n &= n_H + n_T, \\ r &= n_H - n_T = D_n(\omega), \end{aligned}$$

or

$$\begin{bmatrix} n \\ r \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} n_H \\ n_T \end{bmatrix},$$

which is inverted to obtain

$$\begin{bmatrix} n_H \\ n_T \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} n \\ r \end{bmatrix}.$$

For a given value of (n, r), a path from the origin to the point (n, r) exists only if n and r result in positive integer values for n_H and n_T.

In n tosses of the coin, any two outcomes with the *same* number of heads will end up at the same final point (n, r). Thus the number of different paths from the origin to the point (n, r) is

$$N_{n,r} \triangleq \binom{n}{\frac{n+r}{2}} = \binom{n}{n_H}$$

as the $n_H = (n + r)/2$ heads that occur may be in any of $n = n_H + n_T$ places.

Remark For each $\omega \in \Omega_n$, there is a corresponding path given by $(D_1(\omega), D_2(\omega), ..., D_n(\omega)) = (d_1, d_2, ..., d_n)$. There are 2^n paths of length n corresponding to the 2^n outcomes in Ω_n. At time n, the possible values of r are given by

$$r = D_n(\omega) = n_H - (n - n_H) = 2n_H - n.$$

As n_H takes on the values $0, 1, 2, ...n$, it follows that

$$r = D_n(\omega) \in \{-n, -(n-2), ..., n-2, n\}.$$

Example 20 *Number of Paths to* $(4, -2)$
Let $(n, r) = (4, -2)$ so that

$$\begin{bmatrix} n_H \\ n_T \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

and the number of different paths from the origin to $(4, -2)$ is

$$N_{4,-2} = \binom{4}{1} = 4.$$

Explicitly, the four outcomes are (H, T, T, T), (T, H, T, T), (T, T, H, T), and (T, T, T, H) whose paths all end at $(n, r) = (4, -2)$.

Convention
Define $N_{n,r} = 0$ if n and r are not of the form

$$\begin{aligned} n &= n_H + n_T, \\ r &= n_H - n_T \end{aligned}$$

for some nonnegative integers n_H and n_T.
For example, if $n = 4$ and $r = 3$, then

$$\begin{bmatrix} n_H \\ n_T \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 4.5 \\ 0.5 \end{bmatrix}$$

and by our convention $N_{4,3} = 0$.

1.7.1 The Reflection Principle

Let $A = (0, \alpha), B = (n, r)$ with $n > 0, \alpha > 0$, and $r > 0$ be two points in the first quadrant (see Figure 1.4). By the reflection of $A = (0, \alpha)$ with respect to the time axis is meant the point $A' = (0, -\alpha)$. Under these conditions, we can now prove the reflection principle.

Theorem 3 *Reflection Principle*
The number of paths from A to B which touch or cross the t-axis equals the number of all paths from A' to B.

Proof. Consider a path $(d_0 = \alpha, d_1, ..., d_n = r)$ from $A = (0, \alpha)$ to $B = (n, r)$ with $\alpha, r, n > 0$ and having one or more vertices on the t-axis. Let t_1 be the time of the first such vertex as shown in Figure 1.4. In other words, $d_0 = \alpha > 0, d_1 > 0, ..., d_{t_1 - 1} > 0, d_{t_1} = 0$. Now

$$(-d_0 = -\alpha, -d_1, ..., -d_{t_1-1}, d_{t_1} = 0, d_{t_1+1}, ..., d_n = r)$$

is a path from A' to B also having $T = (t_1, 0)$ as its first vertex on the t-axis. The sections AT and $A'T$ are reflections of each other and in this manner a $1-1$ correspondence is set up between the paths from A' to B and the paths from A to B that have a vertex on the t-axis. ■

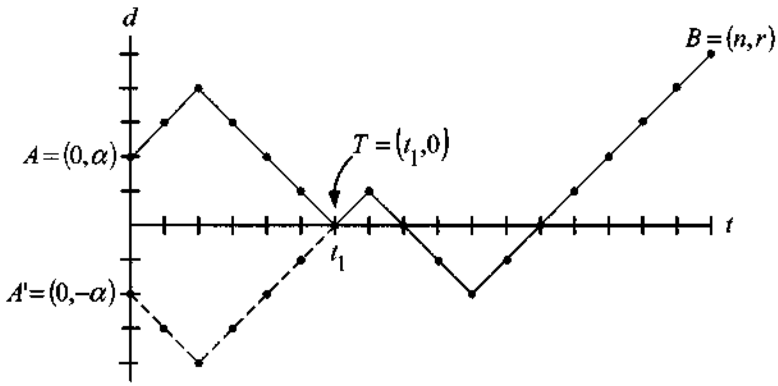

FIGURE 1.4. Illustration of the reflection principle.

Theorem 4 *Ballot Theorem*[2]
Let n and r be *positive* integers. There are precisely

$$N_{n-1,r-1} - N_{n-1,r+1} = \frac{r}{n} N_{n,r} \tag{1.17}$$

paths $(d_0, d_1, ..., d_n = r)$ from the origin $(d_0 = 0)$ to the point (n, r) such that $d_1 > 0, d_2 > 0, ..., d_{n-1} > 0, d_n = r > 0$. We refer to these paths as the *admissible* paths.

Proof. Each of the admissible paths must pass through the point $(1, 1)$. Thus there exist as many of these admissible paths as there are paths from the point $(1, 1)$ to (n, r) which neither touch or cross the t-axis.

(a) $N_{n-1,r-1}$ is the total number of paths from $(1, 1)$ to (n, r).

(b) $N_{n-1,r+1}$ is the total number of paths from the reflected point $(1, -1)$ to (n, r). By the reflection principle this is also the number of paths from $(1, 1)$ to (n, r) which touch or cross the t-axis.

Thus the number of paths the origin $(d_0 = 0)$ to the point (n, r) such that $d_1 > 0, d_2 > 0, ..., d_n = r > 0$ is

$$N_{n-1,r-1} - N_{n-1,r+1}.$$

Now we show $N_{n-1,r-1} - N_{n-1,r+1} = \frac{r}{n} N_{n,r}$. To do so, recall that for a given (n, r) we have

$$\begin{bmatrix} n_H \\ n_T \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} n \\ r \end{bmatrix}.$$

---

[2] See Problem 9 for an interpretation of this theorem to voting.

Therefore, it follows that

$$\begin{aligned}
\begin{bmatrix} n_H - 1 \\ n_T \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} n-1 \\ r-1 \end{bmatrix}, \\
\begin{bmatrix} n_H \\ n_T - 1 \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} n-1 \\ r+1 \end{bmatrix}
\end{aligned}$$

resulting in

$$\begin{aligned}
N_{n-1,r-1} - N_{n-1,r+1} &= \binom{n_H + n_T - 1}{n_H - 1} - \binom{n_H + n_T - 1}{n_H} \\
&= \frac{(n_H + n_T - 1)!}{(n_H - 1)!\, n_T!} - \frac{(n_H + n_T - 1)!}{n_H!\,(n_T - 1)!} \\
&= \frac{n_H}{n_H + n_T}\frac{(n_H + n_T)!}{n_H!\, n_T!} - \frac{n_T}{n_H + n_T}\frac{(n_H + n_T)!}{n_H!\, n_T!} \\
&= \frac{n_H - n_T}{n_H + n_T}\frac{(n_H + n_T)!}{n_H!\, n_T!}.
\end{aligned}$$

As $r = n_H - n_T$ and $n = n_H + n_T$, we have

$$\begin{aligned}
N_{n-1,r-1} - N_{n-1,r+1} &= \frac{r}{n}\binom{n_H + n_T}{n_H} \\
&= \frac{r}{n} N_{n,r}. \qquad (1.18)
\end{aligned}$$

■

Remark Let's check out the expression (1.17) for the extreme case where $r = n$. The total number of paths from $(1,1)$ to (n,n) is given by

$$N_{n-1,r-1}\Big|_{r=n} = N_{n-1,n-1} = \binom{n-1}{\frac{n-1+n-1}{2}} = \binom{n-1}{n-1} = 1,$$

which is clear because this path requires all $n-1$ of the tosses to be heads. On the other hand, with $r = n$, the total number of paths from the reflected point $(1,-1)$ to (n,n) is given by

$$N_{n-1,r+1}\Big|_{r=n} = N_{n-1,n+1} = \binom{n-1}{\frac{n-1+n+1}{2}} = \binom{n-1}{n} = 0.$$

This is simply because in $n-1$ tosses of the coin, one cannot get $n+1$ heads which is what is required to go from $(1,-1)$ to (n,n). This shows that $N_{n-1,n-1} - N_{n-1,n+1} = 1$ and, as $\frac{n}{n}N_{n,n} = 1$, we have a direct validation of (1.17) for $r = n$.

1.8 Random Walks: First Returns and the Main Theorem*

Let's now look at a random walk in which a particle moves n unit steps according to the flip of a coin. Specifically, consider a particle starting at the origin and flip a fair coin at a uniform rate, say every Δt seconds, where the particle moves up or down, respectively, according to whether the outcome of the coin toss is heads or tails, respectively. The position of the particle at time k is denoted as d_k so that $(d_0 = 0, d_1, ..., d_n)$ is the path of the particle.[3] The set of all possible paths at time n is given by $\{(D_1(\omega), D_2(\omega), ..., D_n(\omega))|\ \omega \in \Omega_n\}$ where we recall that

$$D_k(\omega) = \sum_{k=1}^{k} Y_i(\omega) \quad \text{for } k = 1, 2, ..., n.$$

The set $\{\omega \in \Omega_n | D_n(\omega) = r\}$ is the set of all outcomes such that at time n the particle is at the point (n, r). Defining

$$p_{n,r} \triangleq P_n(\{\omega \in \Omega_n | D_n(\omega) = r\}) \tag{1.19}$$

and recalling that

$$N_{n,r} = \binom{n}{n_H} = \binom{n}{\frac{n+r}{2}}$$

is the number paths from the origin $(0,0)$ to the point (n,r), we have

$$p_{n,r} \triangleq P_n(\{\omega \in \Omega_n | D_n(\omega) = r\}) = \binom{n}{\frac{n+r}{2}} 2^{-n}. \tag{1.20}$$

Keep in mind the convention that $\binom{n}{\frac{n+r}{2}}$ is zero unless $\frac{n+r}{2}$ is an integer between 0 and n.

Example 21 $p_{n,r}$

Let $n = 4$ so that for $r \in \{-4, -2, 0, 2, 4\}$ we obtain

$$p_{4,r} \triangleq P_4(\{\omega \in \Omega_4 | D_4(\omega) = r\}) = \binom{4}{\frac{4+r}{2}} 2^{-4},$$

[3] The time k is taken to correspond to the actual time $k\Delta t$.

where $0 \leq n_H = \dfrac{4+r}{2} \leq 4$. We have

$$\begin{aligned}
p_{4,4} &= \binom{4}{4}\frac{1}{2^4} &= \frac{1}{16}, \\
p_{4,2} &= \binom{4}{3}\frac{1}{2^4} &= 4\frac{1}{16}, \\
p_{4,0} &= \binom{4}{2}\frac{1}{2^4} &= 6\frac{1}{16}, \\
p_{4,-2} &= \binom{4}{1}\frac{1}{2^4} &= 4\frac{1}{16}, \\
p_{4,-4} &= \binom{4}{0}\frac{1}{2^4} &= \frac{1}{16}.
\end{aligned}$$

As we have shown, the quantity $\binom{n}{\frac{n+r}{2}}$ is the number of possible paths from the origin to (n, r) in n flips of the coin. Now we ask the question of how many paths go through the same point (n, r) in $n+1$ flips of the coin? Well, after time n, we don't care what happens so ω_{n+1} may be heads or tails or equivalently $Y_{n+1}(\omega) = \pm 1$. Thus the total number of paths from the outcomes in Ω_{n+1} that go through (n, r) is $2\binom{n}{\frac{n+r}{2}}$ and each of these paths has probability $1/2^{n+1}$. Thus

$$\begin{aligned}
P_{n+1}(\{\omega \in \Omega_{n+1} | D_n(\omega) = r\}) &= 2\binom{n}{\frac{n+r}{2}}\frac{1}{2^{n+1}} \\
&= \binom{n}{\frac{n+r}{2}}\frac{1}{2^n} \\
&= P_n(\{\omega \in \Omega_n | D_n(\omega) = r\}).
\end{aligned}$$

Using similar arguments, we have the following theorem.

Theorem 5 *Consistency of Probabilities*

$$P_{n+m}(\{\omega \in \Omega_{n+m} | D_n(\omega) = r\}) = P_n(\{\omega \in \Omega_n | D_n(\omega) = r\}).$$

Proof.

$$\begin{aligned}
P_{n+m}(\{\omega \in \Omega_{n+m} | D_n(\omega) = r\}) &= 2^m\binom{n}{\frac{n+r}{2}}\frac{1}{2^{n+m}} \\
&= \binom{n}{\frac{n+r}{2}}\frac{1}{2^n} \\
&= P_n(\{\omega \in \Omega_n | D_n(\omega) = r\}).
\end{aligned}$$

■

Example 22 *Consistency of Probabilities*

Let $(n,r) = (3,1)$ and $m = 1$. The number of paths in Ω_3 which pass through the point $(3,1)$ is

$$\binom{n}{\frac{n+r}{2}} = \binom{3}{\frac{3+1}{2}} = 3.$$

The number of paths in Ω_4 which go through the point $(3,1)$ is

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

There are two paths in Ω_4 going through $(3,1)$ for each path in Ω_3 that goes through $(3,1)$ as the 4^{th} toss ω_4 can be heads or tails. The six possible paths due to outcomes in Ω_4 which go through $(3,1)$ are shown in Figure 1.5.

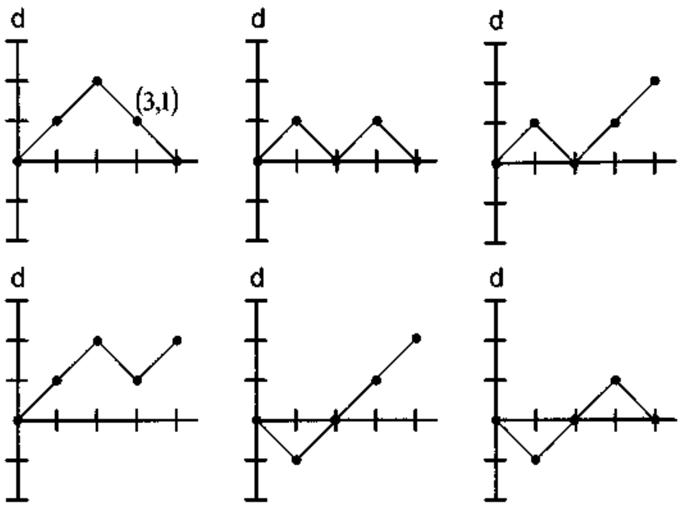

FIGURE 1.5. The six possible paths which go through $(3,1)$.

Each of these six paths has probability $1/2^4$. Then

$$P_4(\{\omega \in \Omega_4 | D_3(\omega) = 1\}) = 2\binom{3}{\frac{3+1}{2}}\frac{1}{2^4} = 6\frac{1}{2^4}$$

while

$$P_3(\{\omega \in \Omega_3 | D_3(\omega) = 1\}) = \binom{3}{\frac{3+1}{2}}\frac{1}{2^3} = 3\frac{1}{2^3}.$$

Definition 5 *Return to the Origin*

A *return to the origin* occurs at time $k \leq n$ for a path corresponding to the outcome $\omega \in \Omega_n$ if $D_k(\omega) = 0$.

Remark Note that k must be *even* for $D_k(\omega) = 0$.

Example 23 *Return to the Origin*

Let $k = 2, n = 4$ so that

$$\begin{aligned} p_{4,0} &\triangleq P_4(\{\omega \in \Omega_4 | D_2(\omega) = 0\}) \\ &= P_2(\{\omega \in \Omega_2 | D_2(\omega) = 0\}) \quad \text{(by Theorem 5)} \\ &= \binom{2}{\frac{2+0}{2}} 2^{-2} \\ &= \frac{1}{2}. \end{aligned}$$

On the other hand, with $k = 3, n = 4$ we have

$$\begin{aligned} p_{4,0} &\triangleq P_4(\{\omega \in \Omega_4 | D_3(\omega) = 0\}) \\ &= P_3(\{\omega \in \Omega_3 | D_3(\omega) = 0\}) \quad \text{(by Theorem 5)} \\ &= \binom{3}{\frac{3+0}{2}} 2^{-3} \\ &= \binom{3}{3/2} 2^{-3} \\ &= 0 \end{aligned}$$

using our convention that $\binom{n}{\frac{n+r}{2}} = 0$ if $\frac{n+r}{2}$ is not a positive integer. This simply because a return to the origin cannot occur at time $k = 3$ as this is an odd number of coin tosses.

As discussed above, at any time $k > 0$ that the particle returns to the origin, there must be an equal number of heads and tails after k tosses of the coin, that is, $k = k_H + k_T = 2k_H$ where k_H is the number of heads in k tosses of the coin and k_T is the number of tails. To study returns to the origin, let's only consider even times for k, i.e., write $k = 2v \leq n$. We already know by (1.20) that the probability of a return to the origin at time $k = 2v$ is given by

$$u_{2v} \triangleq p_{2v,0} = P_{2v}(\{\omega \in \Omega_{2v} | D_{2v}(\omega) = 0\}) = \binom{2v}{v} \frac{1}{2^{2v}},$$

where $\binom{2v}{v}$ is the number of paths from the origin $(0, 0)$ to $(2v, 0)$. Stirling's formula gives us [see (1.9)]

$$u_{2v} \approx \frac{1}{\sqrt{\pi v}},$$

which is a good approximation for large v. For example, with $v = 5$ we have

$$u_{10} = 0.2461, \qquad \frac{1}{\sqrt{\pi v}} = 0.2523$$

while with $v = 10$ we obtain

$$u_{20} = 0.1762, \qquad \frac{1}{\sqrt{\pi v}} = 0.1784.$$

Remark By Theorem 5 with $n \geq 2v$ we have

$$P_n(\{\omega \in \Omega_n | D_{2v}(\omega) = 0\}) = P_{2v}(\{\omega \in \Omega_{2v} | D_{2v}(\omega) = 0\}) = u_{2v}.$$

1.8.1 First Returns

Let the particle move for $2n$ units of time. The particle returns to the origin for the *first time* at time $2v \leq 2n$ for the path corresponding to an outcome ω if $D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2v-1}(\omega) \neq 0$, but $D_{2v}(\omega) = 0$. Define

$$f_{2v} \triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, ..., D_{2v-1}(\omega) \neq 0, D_{2v}(\omega) = 0\}). \quad (1.21)$$

Starting from the origin, the particle can return to zero only after an even number of coin tosses. So the first return time must be an even number, hence the choice $2v$. We will ultimately want to consider the particle returning to the origin at the end of all the coin tosses, hence the final time must be even as well giving the choice $2n$.

Similar to the arguments in Theorem 5 we have

$$\begin{aligned} f_{2v} &\triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, ..., D_{2v-1}(\omega) \neq 0, D_{2v}(\omega) = 0\}) \\ &= P_{2v}(\{\omega \in \Omega_{2v} | D_1(\omega) \neq 0, ..., D_{2v-1}(\omega) \neq 0, D_{2v}(\omega) = 0\}). \end{aligned} \quad (1.22)$$

To see this, let N_{2v} be the number of paths in Ω_{2v} with $D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2v-1}(\omega) \neq 0, D_{2v}(\omega) = 0$, each with probability $1/2^{2v}$. On the other hand, there are $2^{2n-2v} N_{2v}$ paths in Ω_{2n} with $D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2v-1}(\omega) \neq 0, D_{2v}(\omega) = 0$, each having probability $1/2^{2n}$. Thus in the first case, the probability is $N_{2v}/2^{2v}$ while in the second case it is $2^{2n-2v} N_{2v}/2^{2n} = N_{2v}/2^{2v}$.

Example 24 *First Returns*

Let $2n = 4, 2v = 2$ and we compute f_2 and u_2. We have

$$\{\omega \in \Omega_4 | D_2(\omega) = 0\} = \{\omega \in \Omega_4 | D_1(\omega) \neq 0, D_2(\omega) = 0\}$$

as $D_1(\omega)$ can never be zero. Thus

$$\begin{aligned} u_2 &\triangleq P_2(\{\omega \in \Omega_4 | D_2(\omega) = 0\}) = \binom{2}{\frac{2+0}{2}} \frac{1}{2^2} = \frac{1}{2}, \\ f_2 &\triangleq P_4(\{\omega \in \Omega_4 | D_1(\omega) \neq 0, D_2(\omega) = 0\}) = u_2 = \frac{1}{2}. \end{aligned}$$

The eight possible paths corresponding to outcomes in Ω_4 that satisfy $D_2(\omega) = 0$ are shown in Figure 1.6.

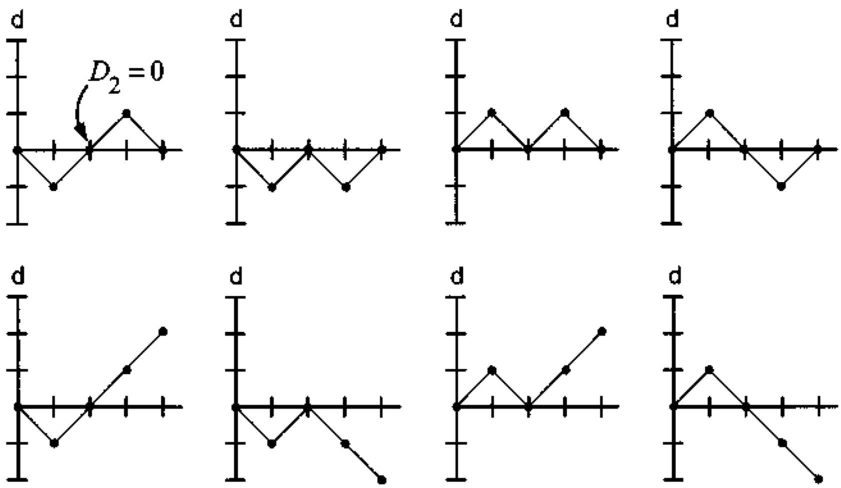

FIGURE 1.6. The eight possible paths with $D_2(\omega) = 0$.

Example 25 *First Returns*

We now compute u_4 and f_4. In this case $v = n = 2$ $(2v = 4, 2n = 4)$. We have

$$\begin{aligned} u_4 &\triangleq P_4(\{\omega \in \Omega_4 | D_4(\omega) = 0\}) = \binom{4}{\frac{4+0}{2}} \frac{1}{2^4} = \frac{6}{16}, \\ f_2 &\triangleq P_4(\{\omega \in \Omega_4 | D_1(\omega) \neq 0, D_2(\omega) \neq 0, D_3(\omega) \neq 0, D_4(\omega) = 0\}) \\ &= P_4(\{\omega \in \Omega_4 | D_2(\omega) \neq 0, D_4(\omega) = 0\}) \\ &= \frac{2}{16} \end{aligned}$$

as there are only two possible paths satisfying $D_2(\omega) \neq 0$ and $D_4(\omega) = 0$ as shown in Figure 1.7.

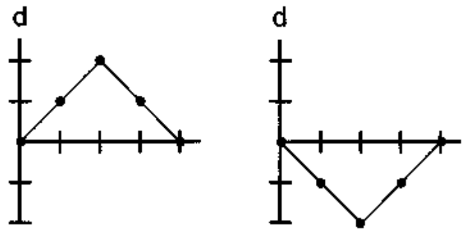

FIGURE 1.7. The two possible paths with $D_2(\omega) \neq 0$ and $D_4(\omega) = 0$.

Convention Define $u_0 \triangleq 1$ and $f_0 \triangleq 0$.
We now show how f_{2n} and u_{2n} are related.

Theorem 6 *Relationship between f_{2n} and u_{2n}*
For $n \geq 1$

$$u_{2n} = f_2 u_{2n-2} + f_4 u_{2n-4} + \cdots + f_{2n} u_0.$$

Proof. A visit to the origin at time $2n$ may be a first return or else the first return occurs at a time $2k < 2n$ and is followed by another return $2n - 2k$ time units later. The probability of this latter case is $f_{2k}u_{2n-2k}$ because there are $2^{2k} f_{2k}$ paths of length $2k$ ending with a first return at time $2k$ and for each of these paths there are $2^{2n-2k}u_{2n-2k}$ paths from the point $(2k, 0)$ to $(2n, 0)$. Thus there are

$$(2^{2k} f_{2k})(2^{2n-2k}u_{2n-2k}) = (f_{2k}u_{2n-2k})2^{2n}$$

paths of length $2n$ with a first return to 0 at time $2k$ that also return to 0 at time $2n$. The first return may occur at any of the times $2, 4, ..., 2k, ..., 2n$ so that the total number of paths with $D_{2n}(\omega) = 0$ is

$$(\textstyle\sum_{k=1}^{n} f_{2k}u_{2n-2k})2^{2n}.$$

Consequently, we have

$$u_{2n} \triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_{2n}(\omega) = 0\}) = \frac{(\sum_{k=1}^{n} f_{2k}u_{2n-2k})2^{2n}}{2^{2n}}$$

or

$$u_{2n} = f_2 u_{2n-2} + f_4 u_{2n-4} + \cdots + f_{2n} u_0.$$

∎

Example 26 *Computation of First Returns*

Using the results of Examples 24 and 25, we have

$$u_6 = f_2u_4 + f_4u_2 + f_6u_0 = \frac{1}{2}\cdot\frac{6}{16} + \frac{2}{16}\cdot\frac{1}{2} + f_6.$$

However,

$$u_6 \triangleq P_6\left(\{\omega \in \Omega_6 | D_6(\omega) = 0\}\right) = \binom{6}{\frac{6+0}{2}}\frac{1}{2^6} = \frac{6!}{3!3!}\frac{1}{2^6} = \frac{5}{16}$$

so that

$$f_6 = \frac{5}{16} - \left(\frac{1}{2}\cdot\frac{6}{16} + \frac{2}{16}\cdot\frac{1}{2}\right) = \frac{1}{16}.$$

The four outcomes resulting in a *first* return to 0 at $t = 6$ are

$$(H,H,H,T,T,T),$$
$$(H,H,T,H,T,T),$$
$$(T,T,T,H,H,H),$$
$$(T,T,H,T,H,H),$$

which is 4 out of $2^6 = 64$ possible outcomes again showing that $f_6 = 4/64 = 1/16$.

1.8.2 Main Theorem

By definition,

$$u_{2n} \triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_{2n}(\omega) = 0\}).$$

The next theorem shows that u_{2n} is also the probability of all outcomes whose corresponding paths *never* go back to 0.

Theorem 7 *Main Theorem*

The probability that *no* return to the origin occurs up to and including time $2n$ is the same as the probability that a return occurs at time $2n$, that is,

$$\begin{aligned} P_{2n}(\{\omega \in \Omega_{2n} | D_1 \neq 0, D_2 \neq 0, ..., D_{2n} \neq 0\}) &= P_{2n}(\{\omega \in \Omega_{2n} | D_{2n} = 0\}) \\ &= u_{2n}. \end{aligned} \tag{1.23}$$

Proof. The event $\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2n}(\omega) \neq 0\}$ occurs only if $D_i(\omega) > 0$ for $i = 1, ..., 2n$ or $D_i(\omega) < 0$ for $i = 1, ..., 2n$, that is,

$$\begin{aligned} \{\omega | D_1 \neq 0, D_2 \neq 0, ..., D_{2n} \neq 0\} = \{&\omega | D_1 > 0, D_2 > 0, ..., D_{2n} > 0\} \cup \\ \{&\omega | D_1 < 0, D_2 < 0, ..., D_{2n} < 0\}. \end{aligned}$$

By symmetry, the two events on the right are equally probable so

$$\begin{aligned}
&P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2n}(\omega) \neq 0\}) \\
&\quad = \quad 2P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}).
\end{aligned} \tag{1.24}$$

As a result of this, we need only prove

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) = \frac{1}{2}u_{2n}.$$

To proceed, recall that $D_{2n}(\omega) \in \{-2n, -2n+2, ..., -2, 0, 2, ..., 2n-2, 2n\}$ (see the remark on page 21). Thus $D_{2n}(\omega) > 0$ implies that $D_{2n}(\omega) \in \{2, ..., 2n-2, 2n\}$. We then have

$$\begin{aligned}
&P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) \\
&\quad = \quad \sum_{r=1}^{n} P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, ..., D_{2n-1}(\omega) > 0, D_{2n}(\omega) = 2r\}).
\end{aligned}$$

By the ballot theorem (Theorem 4), for each $r > 0$, the number of paths satisfying $D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n-1}(\omega) > 0, D_{2n}(\omega) = 2r$ is

$$N_{2n-1,2r-1} - N_{2n-1,2r+1} = \frac{2r}{2n} N_{2n,2r}$$

so that

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1 > 0, ..., D_{2n-1} > 0, D_{2n} = 2r\}) = \frac{N_{2n-1,2r-1} - N_{2n-1,2r+1}}{2^{2n}}.$$

Then

$$\begin{aligned}
&P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) \\
&\quad = \quad \sum_{r=1}^{n} \frac{N_{2n-1,2r-1} - N_{2n-1,2r+1}}{2^{2n}} \\
&\quad = \quad \frac{1}{2^{2n}} \Big((N_{2n-1,1} - N_{2n-1,3}) + (N_{2n-1,3} - N_{2n-1,5}) + \cdots \\
&\qquad\qquad + (N_{2n-1,2n-1} - N_{2n-1,2n+1}) \Big) \\
&\quad = \quad \frac{1}{2^{2n}} (N_{2n-1,1} - N_{2n-1,2n+1}) \\
&\quad = \quad \frac{1}{2^{2n}} N_{2n-1,1}
\end{aligned}$$

since $N_{2n-1,2n+1} = 0$ as explained in the remark after the ballot theorem on page 24. Continuing, we have

$$\frac{1}{2^{2n}} N_{2n-1,1} = \frac{(2n-1)!}{n!\,(n-1)!} \frac{1}{2^{2n}} = \frac{1}{2} \frac{2n!}{n!n!} \frac{1}{2^{2n}} = \frac{1}{2} u_{2n}.$$

That is,

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) = \frac{1}{2}u_{2n} \quad (1.25)$$

so that by (1.24) we have

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2n}(\omega) \neq 0\}) = u_{2n}.$$

■

The next theorem shows that u_{2n} is also the probability that a path never goes negative.

Theorem 8 *Probability of Positive Paths*

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2n}(\omega) \geq 0\}) = u_{2n}.$$

Proof. A path of length $2n$ with all vertices strictly above the t-axis passes through the point $(1, 1)$. Then take the point $(1, 1)$ as a new origin to obtain a path of length $2n - 1$ with all the vertices on or above the new t-axis. We then define

$$N_1 \triangleq \text{Number of outcomes in } \{\omega \in \Omega_{2n} | D_1(\omega) > 0, ..., D_{2n}(\omega) > 0\}$$

which equals

$$\text{Number of outcomes in } \{\omega \in \Omega_{2n-1} | D_1(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0\}.$$

However,

$$\text{Number of outcomes in } \{\omega \in \Omega_{2n-1} | D_1(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0\}$$

equals

$$\frac{1}{2} \underbrace{(\text{Number of outcomes in } \{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0\})}_{\triangleq N_2}$$

Thus

$$\begin{aligned} P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) & \\ = \ & \frac{N_1}{2^{2n}} \\ = \ & \frac{1}{2}\frac{N_2}{2^{2n}} \\ = \ & \frac{1}{2}P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0\}). \end{aligned}$$

However, $D_{2n-1}(\omega)$ cannot be zero as $2n-1$ is an odd number of coin tosses. Thus

$$D_{2n-1}(\omega) \geq 0 \Rightarrow D_{2n-1}(\omega) > 0 \Rightarrow D_{2n}(\omega) \geq 0.$$

That is,

$$\begin{aligned}
&\{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0\} \\
&\quad = \quad \{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0, D_{2n}(\omega) \geq 0\}
\end{aligned}$$

so that

$$\begin{aligned}
&P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) \\
&\quad = \quad \frac{1}{2} P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, ..., D_{2n-1}(\omega) \geq 0, D_{2n}(\omega) \geq 0\}).
\end{aligned}$$

From (1.25) we obtain

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) > 0, D_2(\omega) > 0, ..., D_{2n}(\omega) > 0\}) = u_{2n}/2,$$

from which it follows that

$$P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2n}(\omega) \geq 0\}) = u_{2n}.$$

■

Using the main theorem, we can develop a simple expression for computing the probability of a first return at time $2k$, that is, to compute

$$f_{2k} \triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, ..., D_{2k-1}(\omega) \neq 0, D_{2k}(\omega) = 0\}).$$

Theorem 9 *Expression for First Returns*
For $k = 1, 2, ...n$ we have

$$f_{2k} = u_{2k-2} - u_{2k}.$$

Proof. To proceed, let

$$\begin{aligned}
A &\triangleq \{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2k-2}(\omega) \neq 0\} \\
&= \{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, ..., D_{2k-2}(\omega) \neq 0, D_{2k-1}(\omega) \neq 0\}
\end{aligned}$$

as $2k-1$ is an odd time so $D_{2k-1}(\omega) \neq 0$. Next, define

$$A_2 \triangleq \{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2k}(\omega) \neq 0\}.$$

We want to compute the probability of the set

$$A_1 \triangleq \{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, D_2(\omega) \neq 0, ..., D_{2k-1}(\omega) \neq 0, D_{2k}(\omega) = 0\}.$$

Note that A_1 and A_2 are *disjoint* and

$$A_1 \cup A_2 = A.$$

Thus

$$\begin{aligned} P_{2n}(A_1) &= P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \neq 0, ..., D_{2k-1}(\omega) \neq 0, D_{2k}(\omega) = 0\}) \\ &= P_{2n}(A) - P_{2n}(A_2) \\ &= u_{2k-2} - u_{2k}, \end{aligned}$$

where $P_{2n}(A) = u_{2k-2}$ and $P_{2n}(A_2) = u_{2k}$ follow from the main theorem (Theorem 7). ∎

1.9 Path Maximums and First Visits*

Consider a random walk which starts from the origin $(0,0)$ and goes to some point (n,k). We now investigate the maximum distance it reaches from the origin before ending up at (n,k). This maximum must be greater than or equal to zero because the random walk starts at $D_0(\omega) \equiv 0$. To proceed, the solid line random walk in Figure 1.8 is a path from the origin to the point (n,k) and reaches the line $d = r$. This outcome is in the set

$$\{\omega \in \Omega_n | \max\{D_1(\omega), ..., D_n(\omega)\} \geq r, D_n(\omega) = k\}. \qquad (1.26)$$

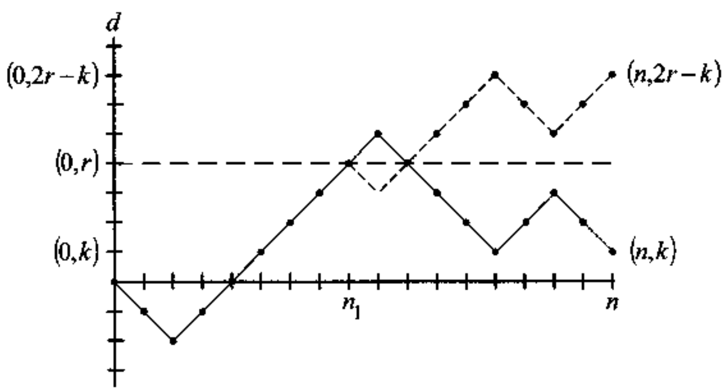

FIGURE 1.8. Reflection of a path about $d = r$.

For each such outcome, there is a unique corresponding path from the origin to the reflected point $(n, 2r-k)$. By reflected point, we simply mean that the point $(n, 2r-k)$ is above the line $d = r$ by the same amount that (n,k) is below $d = r$ at time n. This path is indicated by the dashed line in Figure 1.8. The path to the reflected point is the same as the original path until it intersects $d = r$ at time n_1. After n_1, the remaining part of the path to $(n, 2r-k)$ is obtained by simply taking the results of the tosses from $n_1 + 1$ to n to be the opposite to those of the original (solid line) path. There is a one-to-one correspondence between the paths (outcomes) in (1.26) and the paths (outcomes) in

$$\{\omega \in \Omega_n | D_n(\omega) = 2r - k\}. \tag{1.27}$$

Therefore,

$$\begin{aligned} P_n(\{\omega \in \Omega_n | \max\{D_1, ..., D_n\} \geq r, D_n = k\}) &= P_n(\{\omega \in \Omega_n | D_n = 2r - k\}) \\ &= \binom{n}{\frac{n+2r-k}{2}} \frac{1}{2^n} \\ &= p_{n,2r-k}. \end{aligned}$$

This holds for $k \leq r$. In particular, for $k = r$, we have

$$\{\omega \in \Omega_n | \max\{D_1(\omega), ..., D_n(\omega)\} \geq r, D_n(\omega) = r\} = \{\omega \in \Omega_n | D_n(\omega) = r\}. \tag{1.28}$$

With $k \leq r$, the set of paths that start at $(0,0)$ and end at (n,k) which also have a maximum value *equal* to r are characterized by

$$\{\omega \in \Omega_n | \max\{D_1(\omega), ..., D_n(\omega)\} = r, D_n(\omega) = k\}. \tag{1.29}$$

Then the set of outcomes that start at $(0,0)$ and end at (n,k) with a maximum value greater than or equal to r can be decomposed into two disjoint sets. Specifically, we have

$$\begin{aligned} &\{\omega | \max\{D_1, ..., D_n\} \geq r, D_n = k\} = \\ &\{\omega | \max\{D_1, ..., D_n\} = r, D_n = k\} \bigcup \{\omega | \max\{D_1, ..., D_n\} \geq r + 1, D_n = k\}. \end{aligned}$$

Then computing the probability of both sides gives

$$\begin{aligned} &P_n(\{\omega \in \Omega_n | \max\{D_1(\omega), ..., D_n(\omega)\} \geq r, D_n(\omega) = k\}) = \\ &P_n(\{\max\{D_1, ..., D_n\} = r, D_n = k\}) + P_n(\{\max\{D_1, ..., D_n\} \geq r+1, D_n = k\}). \end{aligned}$$

Finally, rearranging we obtain

$$\begin{aligned} &P_n(\{\omega \in \Omega_n | \max\{D_1(\omega), ..., D_n(\omega)\} = r, D_n(\omega) = k\}) \\ &\quad = P_n(\{\omega | \max\{D_1(\omega), ..., D_n(\omega)\} \geq r, D_n(\omega) = k\}) - \\ &\quad\quad P_n(\{\omega | \max\{D_1(\omega), ..., D_n(\omega)\} \geq r + 1, D_n(\omega) = k\}) \\ &\quad = p_{n,2r-k} - p_{n,2r+2-k}. \end{aligned} \tag{1.30}$$

First Visits to r

We can use the just completed results to compute the probability of a first visit to r. For a random walk to start at $(0,0)$ and get to r for the first time at time n, it must go through the point $(n-1, r-1)$ with $D_k(\omega) \leq r-1$ for $k = 1, ..., n-1$. That is,

$$\{\omega \in \Omega_n | D_1(\omega) < r, ..., D_{n-1}(\omega) < r, D_n(\omega) = r\}$$

$$\begin{aligned} &= \{\omega \in \Omega_n | D_1 \leq r-1, ..., D_{n-2} \leq r-1, D_{n-1} = r-1, D_n = r\} \\ &= \{\omega \in \Omega_n | \max\{D_1, ..., D_{n-1}\} = r-1, D_{n-1} = r-1, D_n = r\} \\ &= \{\omega \in \Omega_n | \max\{D_1, ..., D_{n-1}\} = r-1, D_{n-1} = r-1, Y_n = 1\}. \end{aligned}$$

Therefore

$$\begin{aligned} &P_n(\{\omega \in \Omega_n | D_1(\omega) < r, ..., D_{n-1}(\omega) < r, D_n(\omega) = r\}) \\ &= P_n(\{\omega \in \Omega_n | \max\{D_1, ..., D_{n-1}\} = r-1, D_{n-1} = r-1, Y_n = 1\}) \\ &= P_n(\{\omega \in \Omega_n | \max\{D_1, ..., D_{n-1}\} = r-1, D_{n-1} = r-1\}) P_n(\{\omega | Y_n = 1\}). \end{aligned}$$

Using (1.30) where we substitute $n-1$ for n, $r-1$ for r and set $k = r-1$, the probability of a first visit to r at time n is then given by

$$\begin{aligned} P_n(\{\omega \in \Omega_n | D_1 < r, ..., D_{n-1} < r, D_n = r\}) &= (p_{n-1,r-1} - p_{n-1,r+1})\frac{1}{2} \\ &= (N_{n-1,r-1} - N_{n-1,r+1})\frac{1}{2^{n-1}}\frac{1}{2} \\ &= \frac{r}{n}\binom{n}{\frac{n+r}{2}}\frac{1}{2^n}, \qquad (1.31) \end{aligned}$$

where the last step follows from (1.18) as $n_H = (n+r)/2$.

1.10 Random Walks: Last Visits and Long Leads*

We've been looking at tossing a coin $2n$ times and watching a particle start at 0 and then move up or down according to whether a head or a tail occurred. If the particle goes back to 0, we call this a *visit* or a *return* to the origin. It is also referred to as an *equalization* as the number of heads and the number of tails are equal at this point in time. This random walk can also be viewed as a gambling game. Consider two players who bet a dollar every time the coin is tossed. If heads appear, then player 1 gets a dollar from player 2, while if a tails appears, player 2 gets a dollar from player 1. In this context, an equalization in interpreted as neither player being ahead. The following example is one of the many interesting results which are a consequence of the theory developed above. We quote from W. Feller's book (see [14], page 79)

> Suppose that a great many coin-tossing games (random walks) are conducted simultaneously at the rate of one per second, day and night, for a whole year. On the average, in one of ten games, the last equalization will occur before 9 days have passed, and the lead will not change during the following 356 days. In one out of twenty cases the *last* equalization takes place within 2.5 days, and in one out of a hundred cases it occurs within the first 2 hours and 10 minutes.

The following theorem from W. Feller [14] gives us the tool to verify these remarkable facts.

Theorem 10 *Last Visits to the Origin*
The probability that up to and including time $2n$ the last visit to the origin occurs at time $2k$ is given by

$$\begin{aligned} \alpha_{2k,2n} &\triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_{2k}(\omega) = 0, D_{2k+1}(\omega) \neq 0, ..., D_{2n}(\omega) \neq 0\}) \\ &= u_{2k}u_{2n-2k} \end{aligned} \tag{1.32}$$

for $k = 0, 1, 2, ..., n$.

Proof. We want to compute the probability of the set

$$A \triangleq \{\omega \in \Omega_{2n} | D_{2k}(\omega) = 0, D_{2k+1}(\omega) \neq 0, ..., D_{2n}(\omega) \neq 0\}.$$

The quantity $2^{2k}u_{2k}$ is the number of paths in $2k$ tosses with $D_{2k}(\omega) = 0$. Consider the point $(2k, 0)$ as a new origin. The *main theorem* (Theorem 7) tells us that there are $2^{2n-2k}u_{2n-2k}$ paths for the next $2n - 2k$ time steps that do not touch the t-axis. Thus there are a total of $2^{2k}u_{2k} \cdot 2^{2n-2k}u_{2n-2k}$ paths in A and therefore

$$\begin{aligned} \alpha_{2k,2n} &\triangleq P_{2n}(\{\omega \in \Omega_{2n} | D_{2k}(\omega) = 0, D_{2k+1}(\omega) \neq 0, ..., D_{2n}(\omega) \neq 0\}) \\ &= u_{2k}u_{2n-2k}. \end{aligned}$$

∎

Remark 1 An explicit formula for $\alpha_{2k,2n}$ is simply

$$\alpha_{2k,2n} = u_{2k}u_{2n-2k} = \binom{2k}{k}\binom{2n-2k}{n-k}2^{-2n}.$$

Example 27 $\alpha_{2k,2n}$ for $2n = 4$
We evaluate (1.32) with $2n = 4$ and $k = 0, 1, 2$.
For $k = 0$ it follows that $\alpha_{0,4} = 3/8$ so that there are $2^4(3/8) = 6$ paths of length 4 that never return to 0 after time 0.
For $k = 1$ it follows that $\alpha_{2,4} = 2/8$ so that there are $2^4(2/8) = 4$ paths of length 4 that never return to 0 after time 2.

For $k = 2$ it follows that $\alpha_{4,4} = 3/8$ so that there are $2^4(3/8) = 6$ paths of length 4 that never return to 0 after time 4. This is more clearly understood as the number of paths that return to zero at time 4, i.e., $\alpha_{4,4} = u_4 = P_4(\{\omega \in \Omega_4 | D_4(\omega) = 0\})$.

Example 28 $\alpha_{2k,2n}$ **for** $n = 10$
For $n = 10$ $(2n = 20)$, the following table gives the values of $\alpha_{2k,20}$ for $k = 0, 1, ..., 10$.

	$k=0$ $k=10$	$k=1$ $k=9$	$k=2$ $k=8$	$k=3$ $k=7$	$k=4$ $k=6$	$k=5$
$\alpha_{2k,20}$	0.1762	0.0927	0.0736	0.0655	0.0617	0.0606

1.10.1 Inverse Sine Approximation for $\alpha_{2k,2n}$

For $k < n$ with both k large *and* $n - k$ large, Stirling's approximation [see (1.9)] allows us to write

$$u_{2k} \approx \frac{1}{\sqrt{\pi k}}, \quad u_{2n-2k} \approx \frac{1}{\sqrt{\pi (n-k)}}$$

so that

$$\alpha_{2k,2n} \approx \frac{1}{n} \frac{1}{\pi \sqrt{\frac{k}{n}\left(1 - \frac{k}{n}\right)}}.$$

For $k = 0, 1, ..., n - 1$, define $x_k = k/n$ so that with (see Figure 1.9)

$$f(x) \triangleq \frac{1}{\pi \sqrt{x(1-x)}}$$

we have

$$\alpha_{2k,2n} \approx \frac{1}{n} f(x_k).$$

The probability that the last visit to the origin occurred between times $2k_1$ and $2k_2$ is

$$\sum_{k=k_1}^{k_2} \alpha_{2k,2n}.$$

To compute this sum, we use the above approximation for $\alpha_{2k,2n}$ which is valid as long as $0 << k_1 \leq k \leq k_2 << n$ (so that k_1, k_2, and $n - k_2$ are all large). With $x_1 \triangleq \dfrac{2k_1}{2n}, x_2 \triangleq \dfrac{2k_2}{2n}$ we have

$$\begin{aligned}\sum_{k=k_1}^{k_2} \alpha_{2k,2n} &= \sum_{x_1 \le \frac{k}{n} \le x_2} \alpha_{2k,2n} \\ &\approx \sum_{x_1 \le \frac{k}{n} \le x_2} \frac{1}{\pi} \frac{1}{\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}} \frac{1}{n} && (1.33) \\ &\approx \int_{x_1}^{x_2} \frac{1}{\pi\sqrt{x(1-x)}}\, dx && (1.34) \\ &= \frac{2}{\pi}(\sin^{-1}(\sqrt{x_2}) - \sin^{-1}(\sqrt{x_1})). && (1.35)\end{aligned}$$

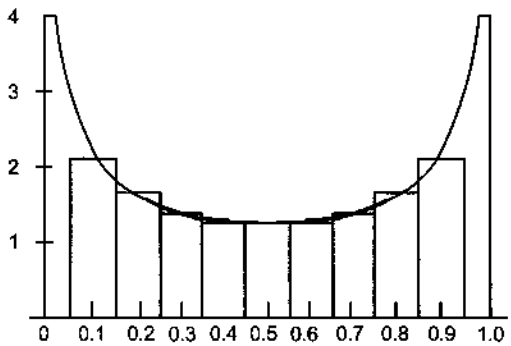

FIGURE 1.9. Graph of $f(x) \triangleq \dfrac{1}{\pi\sqrt{x(1-x)}}$.

Further, we now show that

$$\begin{aligned}\sum_{k=0}^{k_1} \alpha_{2k,2n} &\approx \int_0^{x_1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx, \\ \sum_{k=k_2}^{n} \alpha_{2k,2n} &\approx \int_{x_2}^{1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx\end{aligned}$$

also hold. To proceed, we write

$$\sum_{k=0}^{n} \alpha_{2k,2n} = 1 = \int_0^1 \frac{1}{\pi\sqrt{x(1-x)}}\, dx, \tag{1.36}$$

where the reader is asked to show $\sum_{k=0}^{n} \alpha_{2k,2n} = 1$ in Problem 20 while $\int_0^1 \frac{1}{\pi\sqrt{x(1-x)}}\, dx = 1$ follows from elementary calculus. With $k_2 = n - k_1$, equation (1.36) is then rewritten as

$$\sum_{k=0}^{k_1} \alpha_{2k,2n} + \sum_{k=k_1+1}^{n-k_1-1} \alpha_{2k,2n} + \sum_{k=n-k_1}^{n} \alpha_{2k,2n} =$$
$$\int_0^{x_1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx + \int_{x_1}^{1-x_1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx + \int_{1-x_1}^{1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx.$$

Using the symmetry properties

$$\begin{aligned} f(x) &= \frac{1}{\pi\sqrt{x(1-x)}} = f(1-x), \\ \alpha_{2k,2n} &= \binom{2k}{k}\binom{2n-2k}{n-k}\frac{1}{2^{-2n}} = \alpha_{2n-2k,2n}, \end{aligned}$$

we may rewrite this

$$2\sum_{k=0}^{k_1} \alpha_{2k,2n} + \sum_{k=k_1+1}^{n-k_1-1} \alpha_{2k,2n} = 2\int_0^{x_1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx + \int_{x_1}^{x_2} \frac{1}{\pi\sqrt{x(1-x)}}\, dx,$$

which upon using (1.34) implies[4]

$$\sum_{k=0}^{k_1} \alpha_{2k,2n} \approx \int_0^{x_1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx.$$

It then follows directly that

$$\sum_{k=k_2}^{n} \alpha_{2k,2n} \approx \int_{x_2}^{1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx.$$

Example 29 *Probability of Last Visits to the Origin*

We now verify the examples in the above quote from W. Feller's book [14]. Flipping a coin every second for a year results in

$$2n = 365 \text{ days} \times \frac{24 \text{ hr}}{\text{day}} \times \frac{60 \text{ min}}{\text{hr}} \times \frac{60 \text{ sec}}{\text{min}} = 31{,}536{,}000$$

coin tosses while doing the same thing for 9 days gives

$$2k_1 = 9 \text{ days} \times \frac{24 \text{ hr}}{\text{day}} \times \frac{60 \text{ min}}{\text{hr}} \times \frac{60 \text{ sec}}{\text{min}} = 777{,}600$$

coin tosses. Setting

$$x_1 = \frac{2k_1}{2n} = \frac{9}{365},$$

[4] As $\sum_{k=k_1+1}^{n-k_1-1} \alpha_{2k,2n} \approx \sum_{k=k_1}^{n-k_1} \alpha_{2k,2n} \approx \int_{x_1}^{x_2} \frac{1}{\pi\sqrt{x(1-x)}} dx$.

the probability that the last equalization (last visit to the origin) was in 9 days or less is

$$\begin{aligned}\sum_{k=0}^{k_1} \alpha_{2k,2n} \approx \int_0^{x_1} \frac{1}{\pi\sqrt{x\,(1-x)}}\,dx &= \frac{2}{\pi}\sin^{-1}(\sqrt{x_1}) \\ &= \frac{2}{\pi}\sin^{-1}(\sqrt{9/365}\,) \\ &= 0.10038.\end{aligned}$$

That is, in one out of ten games, the last equalization will occur before nine days have passed!

Similarly, with $x = 2.5/365$, the probability that the last equalization (return to zero) occurred before 2 1/2 days is

$$\frac{2}{\pi}\sin^{-1}(\sqrt{2.5/365}\,) = 0.053 \approx 1/20.$$

Finally, the probability that the last equalization occurred before 2 hours and 10 minutes (which equals 130/60/24 days) is

$$\frac{2}{\pi}\sin^{-1}\left(\sqrt{\frac{130/60/24}{365}}\right) = 0.00997 \approx 1/100.$$

1.10.2 Long Leads

We say a path is on the positive side of the t-axis from the time $2k_1$ to the time $2k_2$ ($k_2 > k_1$) if $D_{2k_1}(\omega) \geq 0, ..., D_{2k_2}(\omega) \geq 0$. For example, the path in Figure 1.10 spends 6 time periods on the positive side, which is from $2k_1 = 6$ to $2k_2 = 8$ and from $2k_3 = 12$ to $2k_4 = 16$ while it spends 10 periods on the negative side.

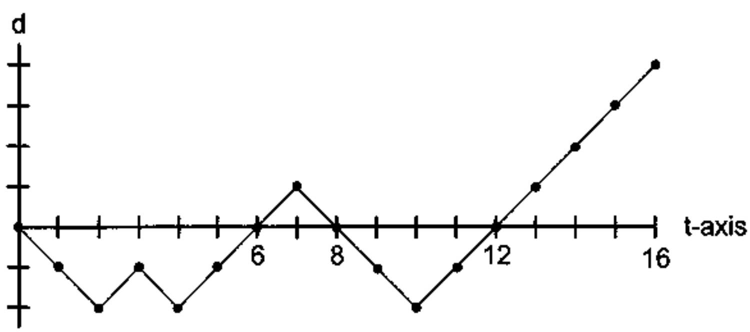

FIGURE 1.10. A path with 6 positive time periods and 10 negative time periods.

Theorem 11 *Inverse Sine Approximation for Positive Leads*

Starting from the origin, the probability that in the time interval from 0 to $2n$, the particle spends $2k$ time periods on the positive side and $2n-2k$ time periods on the negative side is equal to $\alpha_{2k,2n}$.

Further, with $0 < x < 1$, it follows from the inverse sine law approximation that the probability that $x(2n)$ time units or less are spent on the positive side (and thus greater than $(1-x)(2n)$ time periods or more are spent on the negative side) is given by

$$\frac{2}{\pi}\sin^{-1}(\sqrt{x})$$

as $n \to \infty$.

Proof. Consider paths of length $2n$ and let $b_{2k,2n}$ denote the probability that $2k$ periods out of the $2n$ periods are positive, i.e., the path is on the positive side for $2k$ periods. For $k = 0, 1, \ldots, v$ and $v = 1, 2, ..., n$ we need to show that

$$b_{2k,2v} = \alpha_{2k,2v}. \tag{1.37}$$

By Theorem 8,

$$P_{2v}(\{\omega \in \Omega_{2v} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2v}(\omega) \geq 0\}) = u_{2v}$$

which is the probability of $2v$ positive time periods in $2v$ coin tosses. So

$$b_{2v,2v} = u_{2v} = \alpha_{2v,2v}$$

and (1.37) holds for $k = v$. By symmetry, the number of paths for which $2v$ periods are negative in $2v$ tosses has the same probability, that is,

$$b_{0,2v} = u_{2v} = \alpha_{0,2v}.$$

Thus, (1.37) is true for both $k = 0$ and $k = v$ and we now only need show (1.37) holds for $1 \leq k \leq v-1$.

To proceed, let $2k$ of the $2n$ time periods be positive with $1 \leq k \leq v-1$. Then there must be a first return to the origin at some time $2r < 2n$.

Case 1. The first $2r$ time periods up to the *first* return are on the positive side. There are $\frac{1}{2}2^{2r}f_{2r}$ such paths. (There are $2^{2r}f_{2r}$ paths whose first return to 0 is at time $2r$ and half of these are above the t-axis and half are below.) As $0 < r \leq k \leq n-1$ and the path from $(2r, 0)$ to time $2n$ must have exactly $2k-2r$ time periods on the positive side, there are $2^{2n-2r}b_{2k-2r,2n-2r}$ such paths. The total number of these paths is

$$\frac{1}{2}2^{2r}f_{2r}2^{2n-2r}b_{2k-2r,2n-2r}.$$

Then the number of all such paths that have $2k$ of the $2n$ time periods positive *and* their first $2r$ time periods are *above* the t-axis is

$$\sum_{r=1}^{k}\frac{1}{2}2^{2r}f_{2r}2^{2n-2r}b_{2k-2r,2n-2r}. \tag{1.38}$$

Case 2. The first $2r$ time periods up to the first return are spent on the negative side. In this case, after $(2r, 0)$ the path must have $2k$ periods on the positive side requiring $2k \leq 2n - 2r$ so $r \leq n - k$. The number of these paths is

$$\frac{1}{2}2^{2r} f_{2r} 2^{2n-2r} b_{2k,2n-2r}.$$

Then the number of all paths such that $2k$ of the $2n$ time periods are positive *and* their first $2r$ time periods are *below* the t-axis is

$$\sum_{r=1}^{n-k} \frac{1}{2} 2^{2r} f_{2r} 2^{2n-2r} b_{2k,2n-2r}. \tag{1.39}$$

The probability $b_{2k,2n}$ is the sum of (1.38) and (1.39) divided by 2^n, that is,

$$b_{2k,2n} = \sum_{r=1}^{k} \frac{1}{2} f_{2r} b_{2k-2r,2n-2r} + \sum_{r=1}^{n-k} \frac{1}{2} f_{2r} b_{2k,2n-2r}. \tag{1.40}$$

We now use induction on v to prove (1.37). For $v = 1$ we have shown (1.37) is true for $k = 0$. So we set $k = 1$ and consider $b_{2,2}$. This is the probability that in two coin tosses there are two positive periods, which is the same as $P_2(\{\omega \in \Omega_2 | D_1(\omega) \geq 0, D_2(\omega) \geq 0\}) = u_2 = \alpha_{2,2} = \binom{2}{1}\binom{0}{0}2^{-2}$. Thus (1.37) is valid for $v = 1, k = 0, k = 1$. Now assume (1.37) is true for $v \leq n - 1$, that is, by the induction hypothesis we have

$$b_{2k,2v} = \alpha_{2k,2v}$$

for $v \leq n - 1$ and $k = 0, 1, 2, ..., v$. Using this in (1.40) we have

$$\begin{aligned} b_{2k,2n} &= \sum_{r=1}^{k} \frac{1}{2} f_{2r} \alpha_{2k-2r,2n-2r} + \sum_{r=1}^{n-k} \frac{1}{2} f_{2r} \alpha_{2k,2n-2r} \\ &= \frac{1}{2} u_{2n-2k} \sum_{r=1}^{k} f_{2r} u_{2k-2r} + \frac{1}{2} u_{2k} \sum_{r=1}^{n-k} f_{2r} u_{2n-2k-2r}. \end{aligned}$$

By Theorem 6, we know that

$$u_{2n} = f_2 u_{2n-2} + f_4 u_{2n-4} + \cdots + f_{2n} u_0 = \sum_{r=1}^{n} f_{2r} u_{2n-2r}$$

so that

$$\begin{aligned} b_{2k,2n} &= \frac{1}{2} u_{2n-2k} u_{2k} + \frac{1}{2} u_{2k} u_{2n-2k} \\ &= \alpha_{2k,2n}. \end{aligned}$$

Thus (1.37) is true for $v = n$.

With $x_1 = 2k_1/2n$, $0 << k_1 << n$ (i.e., both k_1 and $n - k_1$ are large), the probability that $x_1(2n)$ or less time periods were spent on the positive side is

$$\sum_{k=0}^{k_1} \alpha_{2k,2n} \approx \sum_{0 \le \frac{k}{n} \le x_1} \frac{1}{\pi} \frac{1}{\sqrt{\frac{k}{n}\left(1 - \frac{k}{n}\right)}} \frac{1}{n} \overset{n\to\infty}{\longrightarrow} \int_0^{x_1} \frac{1}{\pi\sqrt{x(1-x)}}\, dx$$

$$= \frac{2}{\pi} \sin^{-1}(\sqrt{x_1}).$$

■

Example 30 *Gambler Leads* [14]

Recall the gambling interpretation of the coin tossing game where two players bet a dollar every time the coin is tossed. If heads appear, then player 1 gets a dollar from player 2, while if a tails appears, player 2 gets a dollar from player 1.

(a) By Theorem 11, the probability that in 20 coin tosses one of the players is always ahead of the other player is equal to

$$\alpha_{0,20} + \alpha_{20,20}.$$

Using the table in Example 28, this is numerically equal to

$$0.1762 + 0.1762 = 0.3522.$$

(b) The probability that in 20 coin tosses that either one of the players is ahead 16 time periods or more is 0.685. To show this, it follows from Theorem 11 that the probability that player 1 is ahead 16 time periods or more is

$$\alpha_{16,20} + \alpha_{18,20} + \alpha_{20,20}$$

and, by the table in Example 28, this is numerically equal to

$$0.0736 + 0.0927 + 0.1762 = 0.3425.$$

The probability that player 2 is ahead 16 time periods or more (equivalently, that player 1 is only ahead 4 times or less) is

$$\alpha_{4,20} + \alpha_{2,20} + \alpha_{0,20} = \alpha_{16,20} + \alpha_{18,20} + \alpha_{20,20} = 0.3425.$$

The set of outcomes for which player 1 is ahead 16 times or more is disjoint from the set of outcomes for which player 2 is ahead 16 times or more. Thus the probability that either player 1 or player 2 is ahead 16 times or more is $2(0.3425) = 0.685$.

Example 31 *Positive Leads and the Inverse Sine Law* [14]

With $2n$ large and $x = 0.024$ (so the number of time periods spent on the positive side are $x(2n)$ or less) it follows by the second part of Theorem 11 that the probability the particle spends greater than $(1-x)(2n) = 0.976(2n)$ time periods on the negative side is asymptotically $(n \to \infty)$ equal to

$$\frac{2}{\pi}\sin^{-1}(\sqrt{0.024}) = 0.099.$$

By symmetry, the probability that the particle spends greater $0.976(2n)$ time periods on the positive side is also 0.099. Adding these two probabilities up, the probability that the particle spends 97.6% of the time on the same side of the t-axis is $2(0.099) = 0.198$ or about 0.2! In words, the experiment is tossing a fair coin a large number of times. If we repeat this experiment many times, then in about 0.2 or 20% of the experiments, the particle will spend 97.6% of its time on just one side of the t-axis.

Example 32 *Positive Leads and the Inverse Sine Law* [14]

Toss a coin every second for 365 days. Again, if heads appear, then player 1 gets a dollar from player 2, while if a tails appears, player 2 gets a dollar from player 1. The probability that player 1 will lead for only 54 days or less is

$$\frac{2}{\pi}\sin^{-1}(\sqrt{54/365}) = 0.251.$$

Thus the probability that either player 1 or player 2 is in the lead for only 54 days or less is twice this value or 0.5.

1.10.3 Experimental Illustration

Figure 1.11 is the output of a computer simulation of a random walk due to 10,000 tosses of a fair coin, that is, a plot of $D_n(\omega)$ versus n. In this particular path, $D_n(\omega)$ is negative for 9562 time steps, positive for 357 time steps, and comes back to zero 81 times. Figure 1.12 is an enlargement of Figure 1.11 for the last 50 time steps, that is, $D_n(\omega)$ versus n for $n = 9950, ..., 10000$. This path where so few time steps are spent one side of the time axis is not so unlikely! To see this, a total of 357 time steps on the positive side gives a fraction of time $357/10000 = 0.0357$ on the positive side. By Theorem 11, the probability that 357 time steps or less are spent on the positive side is given by

$$\frac{2}{\pi}\sin^{-1}(\sqrt{357/10000}) = 0.121.$$

Thus the probability that a particle spends 357 steps or less on *either* the positive side or on the negative side is twice this value, which is 0.242 or about 24%. That is, 24 paths out of 100 will spend 357 time steps or less on one side of the time axis.

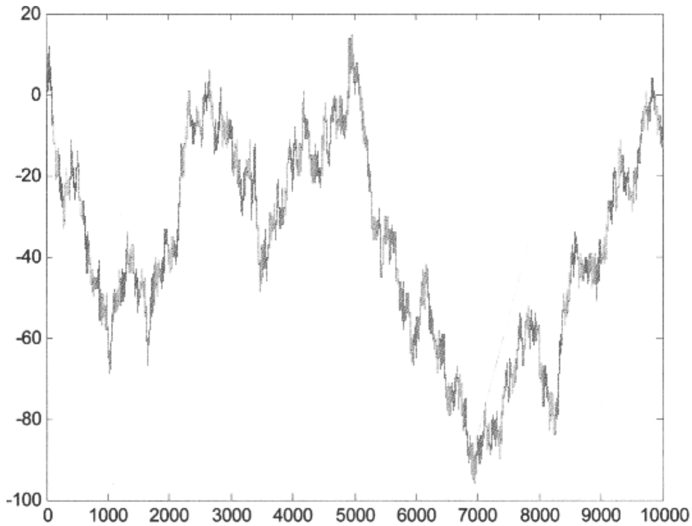

FIGURE 1.11. $D_n(\omega)$ versus n for $n = 1, ..., 10000$.

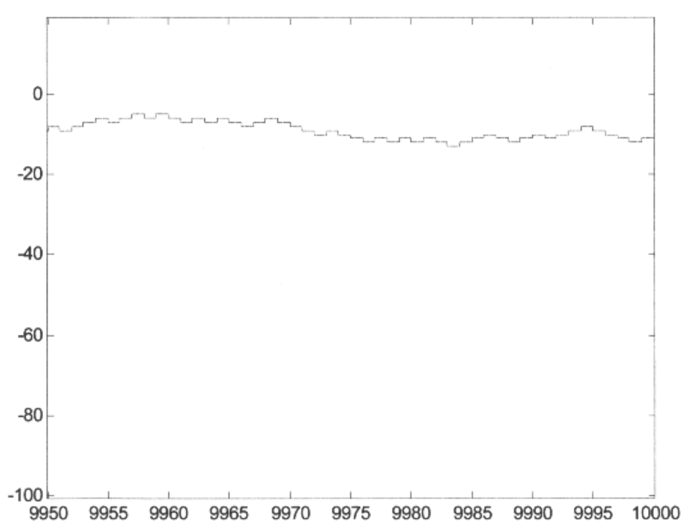

FIGURE 1.12. $D_n(\omega)$ versus n for $n = 9950, ..., 10000$.

The random walk is given by $D_n(\omega) = \sum_{k=1}^{n} Y_k(\omega)$, where the Y_k are independent and $E[Y_k] = 0$ for all k. The (time) average of the first n steps is

$$\frac{D_n(\omega)}{n} = \frac{\sum_{k=1}^{n} Y_k(\omega)}{n}.$$

This is graphed in Figure 1.13 for the path of Figure 1.11. It follows from the *strong law of large numbers* (which will be discussed in later chapters) that

$$\frac{D_n(\omega)}{n} = \frac{\sum_{k=1}^{n} Y_k(\omega)}{n} \xrightarrow{n\to\infty} E[Y_k] = 0.$$

Figure 1.13 does indeed indicate that $D_n(\omega)/n$ goes to 0 as n gets large.

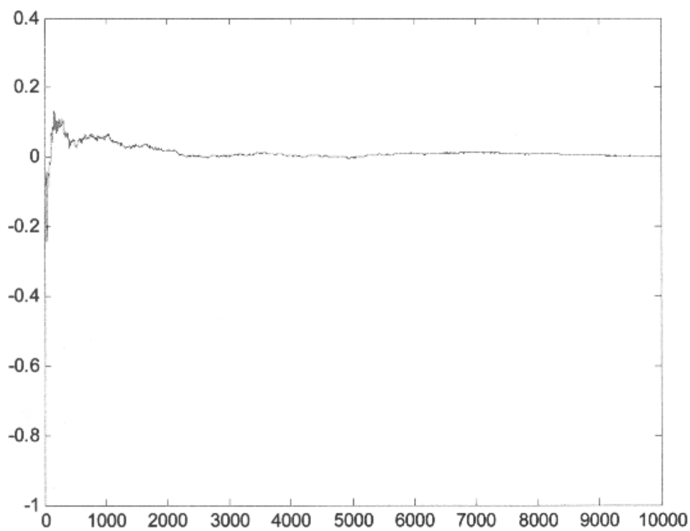

FIGURE 1.13. $\frac{D_n(\omega)}{n} = \frac{\sum_{k=1}^{n} Y_k(\omega)}{n}$ for $n = 1, 2, ..., 10000$.

Appendix: Venn Diagrams

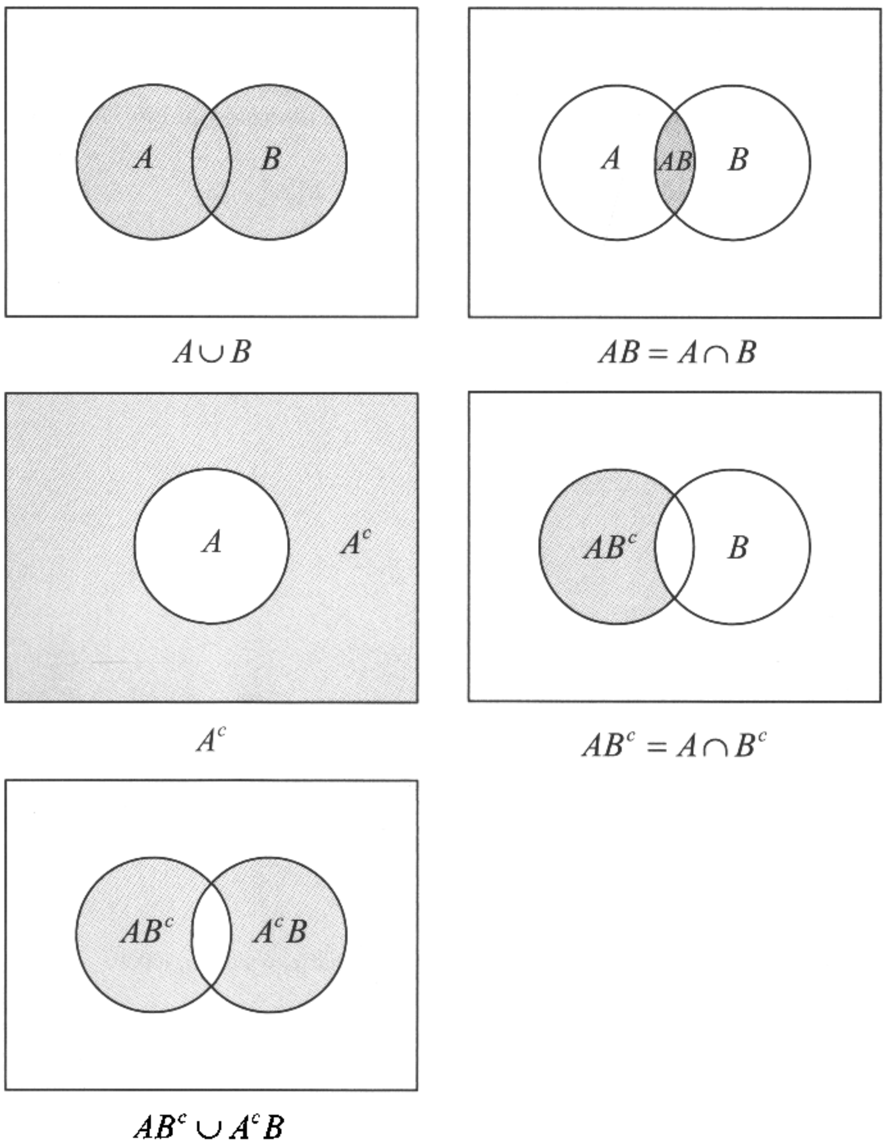

FIGURE 1.14. Venn diagrams illustrating set operations.

Problems

Problem 1 *Expectation*

Let $S_n : \Omega_n \to \mathbb{R}$ be the number of heads in n tosses, i.e., the binomial random variable given in (1.13). Then

$$P_n(\{\omega \in \Omega_n | S_n(\omega) = k\}) = \binom{n}{k} 2^{-n}$$

for $k = 0, 1, ..., n$ and the expected value of S_n is given by

$$E[S_n] \triangleq \sum_{k=0}^{n} k P_n\left(\{\omega \in \Omega_n | S_n(\omega) = k\}\right) = \sum_{k=0}^{n} k \binom{n}{k} 2^{-n}.$$

Show that $E[S_n] = n/2$. Hint: $(sa + b)^n = \sum_{k=0}^{k=n} \binom{n}{k} (sa)^k b^{n-k}$. Take the derivative with respect to s and the set $s = 1, a = 1/2$, and $b = 1/2$.

Problem 2 *Most Likely Value*

Let $S_n : \Omega_n \to \mathbb{R}$ be the number of heads in n tosses, i.e., the binomial random variable given in (1.13). Show that

$$P_n(\{\omega \in \Omega_n | S_n(\omega) = k\}) = \binom{n}{k} 2^{-n}$$

is a maximum for $k = n/2$ if n is even and it is a maximum for $k = (n-1)/2$ or $k = (n+1)/2$ if n is odd. This value of k that gives the maximum probability is called the *most likely value* of S_n.

Hint: First show

$$\frac{P_n(\{\omega \in \Omega_n | S_n(\omega) = k - 1\})}{P_n(\{\omega \in \Omega_n | S_n(\omega) = k\})} = \frac{k}{n - k + 1}$$

and

$$\begin{aligned}
\frac{k}{n-k+1} &< 1 \Leftrightarrow k < \frac{n+1}{2}, \\
\frac{k}{n-k+1} &= 1 \Leftrightarrow k = \frac{n+1}{2} \text{ if } \frac{n+1}{2} \text{ is an integer}, \\
\frac{k}{n-k+1} &> 1 \Leftrightarrow k > \frac{n+1}{2}.
\end{aligned}$$

Problem 3 *Expectation*

Let $X_i, i = 1, ..., n$, be the coin tossing random variables on $(\Omega_n, \mathcal{F}_n, P_n)$. It was shown in Example 15 that $E[X_i] = 1/2$. In a similar fashion, compute the following:

(a) $E[X_1 X_2]$

(b) $E[X_1^2]$

(c) $E[X_i X_j]$ with $i \neq j$

(d) $E[(X_i - 1/2)(X_j - 1/2)]$ with $i \neq j$

(e) $E[(X_i - 1/2)^2]$

Problem 4 *Expectation*

Let $X_i, i = 1, ..., n$, be the fair coin tossing random variables on $(\Omega_n, \mathcal{F}_n, P_n)$. With $S_n(\omega) \triangleq X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)$, compute

$$E[(S_n/n - 1/2)^2].$$

Hint: Note that

$$\frac{S_n}{n} - \frac{1}{2} = \frac{\sum_{i=1}^{n}(X_i - 1/2)}{n}$$

and use the results of Problem 3.

Problem 5 *WLLN versus the Most Likely Value*

The probability of n heads in $2n$ tosses is given by

$$\begin{aligned} P_{2n}(\{\omega \in \Omega_{2n} | \; n \text{ heads in } 2n \text{ tosses}\}) &= P_{2n}(\{\omega \in \Omega_{2n} | \; S_{2n} = n\}) \\ &= \binom{2n}{n}\frac{1}{2^{2n}} \\ &= \frac{(2n)!}{n!n!}2^{-2n}, \end{aligned}$$

In Problem 2 it was shown that $k = n$ gives the maximum value of $\binom{2n}{k}\frac{1}{2^{2n}}$ over $k = 0, 1, ..., 2n$ and, as a result, n is referred to as the *most likely* value of k. It was also shown in Example 6 on Stirling's approximation that

$$\lim_{n\to\infty} P_{2n}(\{\omega \in \Omega_{2n} | \; n \text{ heads in } 2n \text{ tosses}\}) \to \frac{1}{\sqrt{\pi n}} \to 0 \text{ as } n \to \infty.$$

This implies that for any ϵ satisfying $0 < \epsilon < 1$ we have

$$\lim_{n\to\infty} P_{2n}(\{\omega \in \Omega_{2n} | \; |S_{2n} - n| < \epsilon\}) = 0$$

since $\{\omega \in \Omega_{2n} | \; |S_{2n} - n| < \epsilon\}$ is the set of outcomes for which there are *exactly* n heads in $2n$ tosses. Equivalently, we have

$$\lim_{n\to\infty} P_{2n}(\{\omega \in \Omega_{2n} | \; |S_{2n} - n| \geq \epsilon\}) = 1.$$

Replacing n by $2n$ in (1.16), does this contradict the WLLN? Explain why or why not.

Problem 6 *Cumulative Distribution Function of X_i*

For $i = 1, 2, ..., n$ define the cumulative distribution function $F_{X_i}(d)$ of the fair coin tossing random variable X_i as

$$F_{X_i}(d) = P_n(\{\omega \in \Omega_n | \ X_i(\omega) \leq d\}),$$

where $-\infty < d < \infty$.

(a) For $-\infty < d < \infty$, plot $F_{X_i}(d)$.

(b) Note that $F_{X_i}(d)$ is *not* continuous at $d = 0$ and $d = 1$. By direct calculation, show that at these points of discontinuity we have

$$P_n(\{\omega \in \Omega_n | X_i(\omega) = 0\}) = F_{X_1}(0) - F_{X_1}(0^-),$$
$$P_n(\{\omega \in \Omega_n | X_i(\omega) = 1\}) = F_{X_1}(1) - F_{X_1}(1^-).$$

That is, the value of the jumps of $F_{X_i}(d)$ at its discontinuities $d = 0, 1$, respectively, are equal to the probability that X_i equals $0, 1$, respectively. Here

$$F_{X_1}(0^-) \triangleq \lim_{\substack{d \to 0 \\ d < 0}} F_{X_1}(d), \quad F_{X_1}(0^+) \triangleq \lim_{\substack{d \to 0 \\ d > 0}} F_{X_1}(d).$$

Problem 7 *Cumulative Distribution Function of S_n*

Let $X_i, i = 1, ..., n$, be the fair coin tossing random variables on $(\Omega_n, \mathcal{F}_n, P_n)$. With $S_n(\omega) \triangleq X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)$ define the cumulative distribution function $F_{S_n}(d)$ of the binomial random variable S_n as

$$F_{S_n}(d) \triangleq P_n(\{\omega \in \Omega_n | \ S_n(\omega) \leq d\}),$$

for $-\infty < d < \infty$.

(a) Consider S_3 whose range is $\{0, 1, 2, 3\}$. For $-\infty < d < \infty$, plot $F_{S_3}(d)$.

(b) Note that $F_{S_3}(d)$ is *not* continuous at $d = 0, 1, 2, 3$. By direct calculation show for $d = 0, 1, 2, 3$ that

$$P_n(\{\omega \in \Omega_n | S_n(\omega) = d\}) = F_{S_3}(d) - F_{S_3}(d^-).$$

In other words, the value of the jumps of $F_{S_3}(d)$ at its discontinuities $d = 0, 1, 2, 3$, respectively, are equal to the probability that S_3 equals $0, 1, 2, 3$, respectively. Here

$$F_{S_3}(k^-) \triangleq \lim_{\substack{d \to k \\ d < k}} F_{S_3}(d), \quad F_{S_3}(k^+) \triangleq \lim_{\substack{d \to k \\ d > k}} F_{S_3}(d).$$

Remark In general, $F_{S_n}(d)$ is *not* continuous at $d = 0, 1, 2, ..., n$. For $d = 0, 1, 2, ..., n$ we have

$$P_n(\{\omega \in \Omega_n | S_n(\omega) = d\}) = F_{S_n}(d) - F_{S_n}(d^-).$$

In words, the value of the jumps of $F_{S_n}(d)$ at its discontinuities $0, 1, 2, ..., n$, respectively, are equal to the probability that S_n equals $0, 1, 2, ..., n$, respectively.

Problem 8 *Combinatorics: Pascal's Triangle*

(a) Prove Pascal's triangle which states that

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

(b) Use Pascal's triangle to prove by induction that

$$\sum_{m=k}^{k+\ell} \binom{m-1}{k-1} = \binom{k+\ell}{k}.$$

Hint: First show it is true for $\ell = 0$ and $\ell = 1$. Then, assuming it is true for $\ell - 1$, use Pascal's triangle to show it is true for ℓ.

Problem 9 *Combinatorics: Tasting Tea*

There is a famous story about Ronald Fisher (a pioneer in statistics), who was at an afternoon tea party in Cambridge, England in the 1920s. A women was present who claimed she could taste the difference in how a cup of tea with milk was made in the sense of whether the tea was poured in the cup first or if the milk was poured in the cup first.

Consider a way to test this hypothesis using the *hypergeometric probability distribution*. To explain this distribution, suppose you have a group of N objects with K of one type (say type 1) and the remaining $N - K$ of another type (say type 2). Then you choose n of the N objects randomly. By randomly it is simply meant that the objects are mixed together and you cannot tell whether it is type 1 or type 2 before choosing it. The probability of choosing k of K of the type 1 objects and therefore $n-k$ of $N-K$ type 2 objects is

$$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}.$$

Now consider the following tea tasting test. Ten cups of tea are made where five of them have the tea poured in first and then the milk, while the other five have the milk poured in first and then the tea. Then one is asked to taste each cup and decide if the milk was poured first or the tea was poured first. What is the probability that just by chance this person correctly chooses at least 3 of the 5 cups which had the tea poured first? By "chance" just means that (without even tasting the cups of tea) one arbitrarily chooses 5 of them to go into the tea poured first group and thus puts the other five into the milk poured first group.

The book *The Lady Tasting Tea* [16] by David Salsburg is highly recommended for this story and many others on statistics.

Problem 10 *Properties of a Probability Space*

In the text, the probability space $(\Omega_n, \mathcal{F}_n, P_n)$ was defined and shown to have the following three properties:

I. $P_n(\Omega_n) = 1$.

II. $P_n(A) \geq 0$ for all $A \in \mathcal{F}_n$.

III. For $A, B \in \mathcal{F}_n$ with $A \cap B = \emptyset$, then $P_n(A \cup B) = P_n(A) + P_n(B)$.

Using just these three properties (and perhaps the Venn diagrams on page 50 as a guide), prove the following additional properties:

(1) $P(\emptyset) = 0$.

(2) $P(AB^c) = P(A) - P(AB)$.

(3) $P(A^c) = 1 - P(A)$.

(4) $P(A \cup B) = P(A) + P(B) - P(AB)$.

(5) With $B \subset A$, $P(AB^c) = P(A) - P(B)$.

(6) With $B \subset A$, $P(B) \leq P(A)$.

(7) For any $A \in \mathcal{F}_n, 0 \leq P(A) \leq 1$.

Problem 11 *Multinomial Probability Distribution*

Suppose you are given a fat coin so that, after it is tossed, it can land on either of its two faces (disk-shaped) or on its side (cylinder). Let the probability of "heads" be p_1, the probability of "tails" be p_2 and the probability of "sides" be p_3 where $0 < p_i < 1$ and $p_1 + p_2 + p_3 = 1$. We flip this fat coin n times and on each toss there are the three possible outcomes: H, T, or S. Then define

$$\begin{aligned}\Omega_n &\triangleq \{(\omega_1, ..., \omega_n) |\ \omega_i = H, T, \text{ or } S \text{ for } i = 1, ..., n,\\ \mathcal{F}_n &\triangleq \{\text{all possible subsets of } \Omega_n\}.\end{aligned}$$

Then for each outcome $\omega \in \Omega_n$ define

$$P_n(\omega) = P_n(\{(\omega_1, ..., \omega_n)\}) \triangleq p_1^{k_1} p_2^{k_2} p_3^{k_3},$$

where k_1 is the number of "heads" in ω, k_2 is the number "tails" in ω, and k_3 is the number of "sides" in ω. Finally, as usual, for any $A \in \mathcal{F}_n$ define

$$P_n(A) \triangleq \sum_{\omega \in A} P_n(\omega).$$

(a) Show that

I. $P_n(\Omega_n) = 1$.

II. $P_n(A) \geq 0$ for all $A \in \mathcal{F}_n$.

III. For $A, B \in \mathcal{F}_n$ with $A \cap B = \emptyset$, $P_n(A \cup B) = P_n(A) + P_n(B)$.

(b) Define the three random variables S_1, S_2, and S_3 by

$$\begin{aligned} S_1(\omega) &= \text{number of heads in the } n \text{ tosses}, \\ S_2(\omega) &= \text{number of tails in the } n \text{ tosses}, \\ S_3(\omega) &= \text{number of sides in the } n \text{ tosses}. \end{aligned}$$

Then show the probability that we get k_1 heads, k_2 tails, and k_3 sides in n tosses is

$$P(\{\omega | S_1(\omega) = k_1, S_2(\omega) = k_2, S_3(\omega) = k_3\}) = \binom{n}{k_1\, k_2\, k_3} p_1^{k_1} p_2^{k_2} p_3^{k_3}.$$

Problem 12 *Simulation of Bernoulli Random Variables*
In the set of computer programs that accompany this book, run the file `Bernoulli_Distributed_Random_Variables.m`. Try to understand as much of the program as possible. It should become clearer as you progress further in the book.

Problem 13 *Simulation of Bernoulli Random Variables* [3]
In the set of computer programs that accompany this book, run the file `Histogram_for_Bernoulli_Distributed_Random_Variables.m`. This program shows how to make a histogram of the simulation data obtained from Problem 12. Try to understand how histograms are made in this program.

Problem 14 *Simulation of Binomial Random Variables* [3]
In the set of computer programs that accompany this book, run the file `Histogram_for_Binomially_Distributed_Random_Variables.m`.

Problem 15 *Simulation of the Law of Large Numbers*
In the set of computer programs that accompany this book, run the file `Coin_Tossing_and_the_Law_of_Large_of_Numbers.m`. This program is a simulation of the ratio of the number of heads in n tosses to the total number of tosses, i.e., of the law of large numbers.

Problem 16 *Simulation of a Random Walk*
In the set of computer programs that accompany this book, run the file `Random_Walk.m`. This program is a simulation of a random walk.

Problem 17 *Random Walks*

(a) Show that $D_n(\omega) = 0$ if and only if n is an even integer.

(b) Show that if n is an even integer, $r = D_n(\omega)$ must be an even integer.

(c) Show that if n is an odd integer, $r = D_n(\omega)$ must be an odd integer.

Problem 18 *Ballot Theorem*

Prove the ballot theorem using the fact that all admissible paths (i.e., those paths $(d_0, d_1, ..., d_n = r)$ from the origin $(d_0 = 0)$ to the point (n, r) such that $d_1 > 0, d_2 > 0, ..., d_{n-1} > 0, d_n = r > 0)$ must pass through the point $(2, 2)$.

Problem 19 *Interpretation of the Ballot Theorem*

Suppose we have two candidates for public office. We assume that John Q. Public votes by tossing a fair coin. If heads appears, he votes for candidate A otherwise he votes for candidate B. Let $n = n_H + n_T$ be the total number of people who vote. Given that candidate A wins with exactly n_H votes where $n_H > n_T$, use the ballot theorem to show that the probability he leads throughout the voting is $(n_H - n_T) / (n_H + n_T)$.

Problem 20 *Last Visits to the Origin*

Using the *last visits to the origin* theorem (Theorem 10), show that

$$\sum_{k=0}^{n} \alpha_{2k,2n} = \sum_{k=0}^{n} u_{2k,2n-2k} = 1.$$

Hint: No calculations are needed!

Problem 21 *Probability of Paths that Do Not Cross the t-axis*

By Theorem 8 and symmetry, we have

$$\begin{aligned} u_{2n} &= P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \geq 0, D_2(\omega) \geq 0, ..., D_{2n}(\omega) \geq 0\}) \\ &= P_{2n}(\{\omega \in \Omega_{2n} | D_1(\omega) \leq 0, D_2(\omega) \leq 0, ..., D_{2n}(\omega) \leq 0\}). \end{aligned}$$

Show that the probability that the particle crosses the t-axis in $2n$ time steps is $1 - 2u_{2n}$ and that this goes to 1 as $n \to \infty$.

Problem 22 *The Probability of Returning to the Origin*

Use the results of Theorem 9 to show

$$\sum_{k=1}^{n} f_{2k} = 1 - u_{2n}.$$

Next, using Stirling's approximation, show that

$$\sum_{k=1}^{n} f_{2k} \to 1 \quad \text{as } n \to \infty.$$

Finally, show that

$$f_{2n} = \frac{1}{2n-1} u_{2n}.$$

Problem 23 *Last Visits to the Origin*

In Theorem 10 it was shown that the probability that the last return to the origin is at time $2k$ in $2n$ time steps (coin tosses) is

$$\alpha_{2k,2n} = u_{2k}u_{2n-2k} = \binom{2k}{k}\binom{2n-2k}{n-k}2^{-2n}.$$

Let $n = 2$ $(2n = 4)$.

(a) With $k = 0$ $(2k = 0)$ we have

$$\alpha_{0,4} = \binom{0}{0}\binom{4}{2}\frac{1}{2^4} = \frac{6}{16}$$

so that 6 out of the possible 16 paths have their last visit to the origin at time 0. Draw these 6 paths.

(b) With $k = 1$ $(2k = 2)$ we have

$$\alpha_{2,4} = \binom{2}{1}\binom{2}{1}\frac{1}{2^4} = \frac{4}{16}$$

so that 4 out of the possible 16 paths have their last visit to the origin at time 2. Draw these 4 paths.

(c) With $k = 2$ $(2k = 4)$ we have

$$\alpha_{0,4} = \binom{4}{2}\binom{0}{0}\frac{1}{2^4} = \frac{6}{16}$$

so that 6 out of the possible 16 paths have their last visit to the origin at time 4. Draw these 6 paths.

Problem 24 *Sojourn Times*

In Theorem 11 it was shown that starting from the origin, the probability that in the time interval from 0 to $2n$, the particle spends $2k$ time periods on the positive side and $2n - 2k$ time periods on the negative side is equal to

$$\alpha_{2k,2n} = u_{2k}u_{2n-2k} = \binom{2k}{k}\binom{2n-2k}{n-k}2^{-2n}.$$

Let $n = 2$ $(2n = 4)$.

(a) With $k = 0$ $(2k = 0)$ we have

$$\alpha_{0,4} = \binom{0}{0}\binom{4}{2}\frac{1}{2^4} = \frac{6}{16}$$

so that 6 out of the possible 16 paths spend no time units on the positive side and 4 time units on the negative side. Draw these 6 paths.

(b) With $k = 1$ $(2k = 2)$ we have

$$\alpha_{2,4} = \binom{2}{1}\binom{2}{1}\frac{1}{2^4} = \frac{4}{16}$$

so that 4 out of the possible 16 paths spend 2 time units on the positive side and 2 time units on the negative side. Draw these 4 paths.

(c) With $k = 2$ $(2k = 4)$ we have

$$\alpha_{0,4} = \binom{4}{2}\binom{0}{0}\frac{1}{2^4} = \frac{6}{16}$$

so that 6 out of the possible 16 paths spend 4 time units on the negative side and no time units on the positive side. Draw these 6 paths.