A random walk can be compared to an ideal coin tossing game in which the probability of getting a head or a tail is absolutely random. We can record the successive cumulative gains as the partial sums $S_1, S_2, \ldots, S_k$ of $k$ successive coin tossings as the path of a random walk. For geometric description it is convenient to assume that the tossings are performed at a uniform rate so that the $n^{th}$ trial occurs at epoch $n$. Epoch is used to denote points on the time axis. Mathematically, the word time will refer to an interval or duration. A physical experiment may take some time, but our ideal trials are timeless and occur at epochs.

The successive partial sums $S_1, \ldots, S_n$ will be marked as points on the vertical $x$-axis. These points are called positions of a particle performing a random walk. The path $P_0$ to $P_8$ in the Figure 3.1 represents a random walk of eight steps which returns to the origin at the 8th step.
Figure 3.1. Illustrating the positive paths in a one-dimensional random walk.

In a random walk experiment, each path of length \( p \) can be considered as the outcome of a random walk experiment. For a random walk with \( p \) steps, the probability of each path is \( 2^{-p} \), since there are \( 2^p \) possible paths. This type of random walk is called symmetric random walk.

Let us consider that a particle starts from the initial position \( z \) and moves at regular time intervals, a unit step in the positive or negative direction depending on whether the corresponding trial resulted in the probabilities of success and failure, respectively. The trials terminate when the particle, for the
first time, reaches either 0 or a. Then the particle is said to execute a random walk with absorbing barriers at 0 and a. This random walk is restricted to the possible positions, 1, 2, . . . , a-1. In the absence of absorbing barriers the random walk is called unrestricted. In this case p = q. But in the case of symmetric random walk p = q = $\frac{1}{2}$.

Consider a random walk with \( p \) steps, for any \( k \leq p \), if we arbitrarily fix the 1st \( k \) steps, exactly \( 2^{p-k} \) paths will satisfy these \( k \) conditions. Thus it is concluded that an event determined by the first \( k \leq p \) steps has a probability independent of \( p \). Therefore the total number of steps plays no role if it is sufficiently large. In other words, any path of length \( n \) can be taken as the initial section of a very long path and there is no need to specify the latter length.

In the general theory the notations \( X_1, X_2, \ldots \) generally represent the individual steps and \( S_1, S_2, \ldots \) represent the positions of the particle. Thus

\[
S_n = X_1 + X_2 + \ldots + X_n \tag{3.1}
\]

\[
S_0 = 0 \tag{3.2}
\]

The event “at epoch \( n \), the particle is at the point \( r \)” will be denoted by \( \{S_n=r\} \). Its probability is written as \( P_{n,r} \). This event can also be stated as a visit to \( r \) at the epoch \( n \).

A path from the origin to an arbitrary point \( (n,r) \) exists only if \( n \) and \( r \) are of the form

\[ n = p + q \text{ and } r = p - q \]

where \( p \) and \( q \) are the number of forward and backward steps respectively. The number \( N_{n,r} \) paths from the origin to the point \( (n,r) \) is given by

\[
N_{n,r} = \binom{p+q}{p} = \binom{n}{\frac{n+r}{2}}
\]
and hence the probability for a visit to \( r \) at the epoch \( n \) is given by

\[
P_{n,r} = P \{ S_n = r \} = \binom{n}{(n+r)/2} 2^{-n}
\]  

(3.4)

### 3.1 Returns to origin

A return to origin occurs at epoch \( k \) if \( S_k = 0 \), where \( k \) is necessarily even, i.e., \( k = 2v \). The probability of a return to origin at epoch \( 2v \) is represented as \( P_{2v,0} \). This probability is usually denoted by \( U_{2v} \).

\[
U_{2v} = \binom{2v}{v} 2^{-2v}
\]

(3.5)

A first return occurs at epoch \( 2v \) if

\[
S_1 \neq 0, S_2 \neq 0, \ldots, S_{2v-1} \neq 0 \text{ but } S_{2v} = 0
\]

The probability for this event will be denoted by \( f_{2v} \). By definition \( f_0 = 0 \).

### 3.2 Relation between \( f_{2n} \) and \( U_{2n} \)

A visit to the origin at epoch \( 2n \) may be the first return or the first return occurs at an epoch \( 2k < 2n \) and is followed by a further return to origin at \( 2n \) i.e., \( 2n-2k \) time units later. The probability of such a case is \( f_{2k} U_{2n-2k} \), because there are \( 2^{2k} f_{2k} \) paths of length \( 2k \) ending with a first return, and \( 2^{2n-2k} \times U_{2n-2k} \) paths from the point \((2k,0)\) to \((2n,0)\). Thus we can write

\[
U_{2n} = f_{2} U_{2n-2} + f_{4} U_{2n-4} + \ldots + f_{2n} U_{0} \quad n \geq 1
\]

(3.6)
3.3 **Theory of one dimensional random walk**

A path of \( n \) steps can be considered as an ideal coin tossing game consisting of \( n \) successive tosses of a coin. Let +1 stands for head and -1 for tail. \( S_k \) is the sum of +1 and -1 at the \( k^{th} \) trial. By repeating this trial again for \( n \) tossings we get either same or different values for \( S_k \) at the \( k^{th} \) trial. This represents another path. The probability that a particle is at position \( r \) after \( n \) steps is given by

\[
P(S_n=r) = P_{n,r} = \frac{n! \left(\frac{1}{2}\right)^n}{\left[\frac{(n-r)}{2}\right]! \left[\frac{(n+r)}{2}\right]!}
\]  

We now consider two important results in the theory of the one dimensional random walk.

3.3.1 **Reflection principle**

In one dimensional random walk each step can assume either +1 or -1 value. Let there be \( p \) plus ones and \( q \) minus ones, so that \( p+q=n \). The partial sum \( S_k \) is the difference between the number of plus ones and minus ones occurring at the first \( k \) places. Then,

\[S_kS_{k+1} = \pm 1, \quad S_0 = 0, \quad S_n = p-q \]

where \( k \) can be 1, 2, ..., \( n \).

One dimensional random walk can be represented geometrically in a rectangular coordinate system. The \( t \)-axis is referred to as horizontal axis and \( x \)-axis vertical. The points on the time axis is usually referred to as epochs.

The outcomes ±1 is represented by a polygonal line whose \( k^{th} \) side has slope ±1 and whose \( k^{th} \) vertex has ordinate \( S_k \). Such lines are called paths. This is illustrated in Figure 3.1.
3.3.2 Statement of the reflection principle

The total number of paths from a point $P$ to $Q$ above the $t$-axis that either touch or cross the $t$-axis is equal to the total number of paths from the reflection of $P$ about the $t$-axis (i.e., $P'$) to $Q$.

Consider a path from $P$ to $Q$ in Figure 3.2, which has one or more vertices on the $t$-axis. Let $T$ be the abscissa of the 1st vertex, i.e., up to the value of $T$, the path does not touch or cross the $t$-axis. Then each path from $P$ to $T$ will have its reflection on the $t$-axis. So the number of paths from $P$ to $T$ and $P'$ to $T$ are the same. Therefore, there is a one to one correspondence for all paths from $P$ to $Q$ that touches or crosses the $t$-axis and that of $P'$ to $Q$. This proves the reflection principle.

Figure 3.2. Illustration of the reflection principle
An example of the use of the reflection principle is the Ballot theorem.

3.3.3 The Ballot theorem

Suppose there are two candidates in an election and at the end of the election A has got ‘a’ votes and B has got ‘b’ votes; then the Ballot theorem states that in the counting process, the probability that A is always leading is \( \frac{(a-b)}{(a+b)} \).

According to reflection principle, which is illustrated in Figure 3.3 that number of paths from A to B that cross or touch the t-axis = the total number of paths from A' to B.

![Illustration of the reflection principle.](image-url)
The probability that A is always leading is the same as the probability of the paths from A to B that do not touch or cross the t-axis. This is given by the ratio of the number of paths that do not touch or cross the t-axis to the total number of paths from (0,0).

For A to be always leading the first vote should be for A. Then in the counting process the probability that A is always leading is obtained by counting the number of paths from (1,1) in Figure 3.4 that do not touch or cross the t-axis.

![Diagram](image)

**Figure 3.4** New coordinate system with (1,1) as the new origin.
This is equal to the total number of paths from (1,1)—the number of paths that touch or cross the t-axis

\[ N_{(n-1),(r-1)} - N_{(n-1),(r+1)} \]

\[ = \frac{(n-1)!}{[(n-r)/2]! [(n+r)/2 \cdot 1]!} - \frac{(n-1)!}{[(n-r)/2 - 1]! [(n+r)/2]!} \]

\[ = \frac{(n-1)!}{[(n-r)/2 - 1]! [(n+r)/2 - 1]!} \times \frac{4r}{(n-r)(n+r)} = \frac{(n-1)! r}{[(n-r)/2]! [(n+r)/2]!} \]

\[ \therefore \text{ The probability that } A \text{ is always leading } = \frac{1}{n!} \frac{[(n-r)/2]! [(n+r)/2]!}{[(n-r)/2]! [(n+r)/2]!} = r/n = [(a-b)/(a+b)] \]

3.3.4 The main Lemma

A return to origin occurs at epoch \( k \) if \( S_k = 0 \). Here the value of \( k \) should always be even or \( k = 2n \) and the probability of a return to origin is written as \( P_{2n,0} = U_{2n} \). In the case of one dimensional random walk there are two possibilities for a single step. The condition for a return to origin is that, there must be equal number of forward and backward steps. For return to origin occurs at the \( 2n^{th} \) step, there must be \( n \) forward and \( n \) backward steps.

\[ \therefore \text{ The total number of paths that return to origin is } 2^n C_n \]

Therefore \( U_{2n} = 1/2^{2n} \times 2^n C_n \)

\[ (3.9) \]
Statement

The Main Lemma states that the probability that no return to the origin occurs up to and including epoch $2n$ is the same as the probability that a return occurs at epoch $2n$. This can be mathematically stated as

$$P(S_1 \neq 0, ..., S_{2n} \neq 0) = P(S_{2n} = 0) = U_{2n} \quad (3.10)$$

Proof

Let the probability of a return to origin be $U_{2n}$.

$P(S_1 \neq 0, S_2 \neq 0, ... S_{2n} \neq 0)$ means either all the $S_k$'s be $+$ve or all $-$ve.

So we can rewrite it as

$$P(S_1 > 0, S_2 > 0, ... S_{2n} > 0) = \frac{1}{2} U_{2n} \quad (3.10a)$$

Considering all the possible values of $S_{2n}$.

$$P(S_1 > 0, S_2 > 0, ... S_{2n} > 0)$$

$$= \sum_{r=1}^{2n-1} P(S_1 > 0, S_2 > 0, ... S_{2n-1} > 0, S_{2n} = 2r) \quad (3.10b)$$

all the terms with $r > n$ vanish.

According to the Ballot theorem R. H. S. of equation 3.10b will be

$$\sum_{r=1}^{n} \left(\frac{N_{2n-1, 2r-1} - N_{2n-1, 2r+1}}{2^{2n-1}}\right) \times \frac{1}{2} \quad (3.10c)$$

$$= \sum_{r=1}^{n} \left[\frac{N_{2n-1, 2r-1}}{2^{2n-1}} - \frac{N_{2n-1, 2r+1}}{2^{2n-1}}\right] \times \frac{1}{2},$$

$$= \frac{1}{2} \sum_{r=1}^{n} (P_{2n-1, 2r-1} - P_{2n-1, 2r+1})$$

Since

$$\frac{N_{2n-1, 2r-1}}{2^{2n-1}} = P_{2n-1, 2r-1}$$
On expansion of the terms in the bracket we get

\[ \frac{1}{2} \left[ P_{2n-1,1} - P_{2n-1,3} + P_{2n-1,3} - P_{2n-1,5} + P_{2n-1,5} \ldots + P_{2n-1,2n-1} - P_{2n-1,2n+1} \right] = \frac{1}{2}(P_{2n-1,1} - P_{2n-1,2n+1}) \]

But \( P_{2n-1,2n+1} = 0 \)

\[ \therefore \text{R.H.S.} = \frac{P_{2n-1,1}}{2} \quad (3.10d) \]

\[ P_{2n-1,1} = \binom{2n-1}{2n-1} \cdot 2^{(2n-1)} = \binom{2n-1}{n} \cdot 2^{(2n-1)} \]

\[ P_{2n-1,1} = \frac{(2n-1)!}{n!(n-1)!} 2^{(2n-1)} \quad (3.10e) \]

\[ U_{2n} = P_{2n,0} = \binom{2n}{n} \cdot 2^{2n} = \frac{2n!}{n! \cdot n!} \cdot 2^{2n} = \frac{2n(2n-1)! \cdot 2^{(2n-1)}}{n! \cdot n(n-1)! \cdot 2} \]

On simplification we get

\[ U_{2n} = P_{2n,0} = \frac{(2n-1)!}{n!(n-1)!} 2^{(2n-1)} \quad (3.10f) \]

Equations 3.10e and 3.10f are equal.

\[ \therefore P_{2n-1,1} = P_{2n,0} = U_{2n} \quad (3.10g) \]

The Main Lemma can be stated in a different form.

\[ \therefore P(S_1 \geq 0, \ldots S_{2n} \geq 0) = U_{2n} \quad (3.10h) \]

A path of length \( 2n \) with all vertices above the x-axis pass through the point (1,1). Taking this as a new origin we get a path of length \( (2n-1) \) with all vertices above or on the new axis.
Number of paths in both cases is the same, but the number of steps is $(2n-1)$

\[
\sum_{i=1}^{n} \frac{(N_{2n-1,2r-1} - N_{2n-1,2r+1})}{2^{2n-1}} = \sum_{i=1}^{n} (P_{2n-1,2r-1} - P_{2n-1,2r+1})
\]

\[
= P_{2n-1,1} = U_{2n}
\]

$S_{2n-1} \geq 0$ implies that $S_{2n} \geq 0$ since each outcome can have values either $+1$ or $-1$. 

Figure 3.5. New Co-ordinate system with $(1,1)$ as the new origin.
Hence \( P(S_1 \geq 0, S_2 \geq 0, \ldots, S_{2n} \geq 0) = P(S_1 \geq 0, S_2 \geq 0, \ldots, S_{2n-1} \geq 0) \) \hspace{1cm} (3.10i)

But \( P(S_1 > 0, S_2 > 0, \ldots, S_{2n} > 0) = \frac{1}{2} P(S_1 \geq 0, S_2 \geq 0, \ldots, S_{2n} \geq 0) \) \hspace{1cm} (3.10j)

\[ P(S_1 > 0, S_2 > 0, \ldots, S_{2n} > 0) = \frac{1}{2} U_{2n} \] \hspace{1cm} (3.10k)

Comparing equations 3.10j and 3.10k we get

\[ P(S_1 \geq 0, S_2 \geq 0, \ldots, S_{2n-1} \geq 0) = U_{2n} \] \hspace{1cm} (3.10l)

From equations 3.10i and 3.10l we have \[ P(S_1 \geq 0, S_2 \geq 0, \ldots, S_{2n} \geq 0) = U_{2n} \] \hspace{1cm} (3.10)

3.3.5 Probability of first return to origin—\( F_{2n} \)

The probability of first return to origin can be mathematically defined as \( P(S_1 \neq 0, S_2 \neq 0, \ldots, S_{2k} \neq 0) \) are satisfied for \( k=n-1 \) but not for \( k=n \), i.e., \( S_{2k}=0 \) for \( k=n \).

According to main Lemma,

\[ F_{2n} = U_{2n-2} - U_{2n} \text{ where } n = 1, 2, \ldots \] \hspace{1cm} (3.11)

Total number of paths of length 2n with no return to origin occurs up to \((2n-2) = U_{2n-2} \times 2^{2n} \). \( U_{2n-1} \) cannot occur since \((2n-1) \) is an odd number and odd number of steps cannot return to origin.

\[ \therefore \text{ The total number of paths with a first return to origin at } 2n \text{ is given by the equation.} \]

\[ F_{2n} = \frac{2^{2n}(U_{2n-2} - U_{2n})}{2^{2n}} = \frac{U_{2n-2} - U_{2n}}{2^{2n}} = \frac{2n-2}{n-1} \cdot 2^{(2n-2)} - \binom{2n}{n} 2^{2n} \]
On simplification we get,

\[
\begin{align*}
= \frac{(2n-2)! 2^{(2n-2)}}{(n-1)! (n-1)!} & - \frac{2n! 2^{2n}}{n! n!} = \frac{2n! 2^{2n}}{n! n!} \left[ \frac{n \times n \times 2^2}{(2n-1)2n} - 1 \right] \\
= \frac{2n! 2^{2n}}{n! n!} \times \frac{2n}{2n-1} - 1
\end{align*}
\]

On simplification we get,

\[
F_{2n} = \frac{U_{2n}}{(2n-1)} \quad (3.12)
\]

### 3.4 Multidimensional random walk

In two dimensional random walk, the particle moves along the x and y coordinates in unit steps. So there are 4 possibilities for each step and the probability for each step is \( \frac{1}{4} \). In a symmetric random walk all the directions have the same probability. The probability of a return to origin decreases as the dimensionality increases. Polya stated an interesting theorem concerning a return to origin in multidimensional random walk.

**Theorem**

In symmetric random walks in one and two dimensions there is probability one that the particle will infinitely often return to its initial position. In three dimensions, however, this probability is only about 0.35.

In two dimensions a return to origin is possible only if the number of steps in the positive x and y directions is equal to those in the negative x and y directions, respectively. Hence \( U_n = 0 \) if \( n \) is odd.
Let \( k \) steps be taken in the positive x-direction and \((n-k)\) steps in the positive y-direction. Then a return to origin \( U_{2n} \) is given by the equation.

\[
U_{2n} = 1/4^{2n} \sum_{k=0}^{n} \frac{2n!}{k! \cdot k! \cdot (n-k)! \cdot (n-k)!}
\]

\[
= 1/4^{2n} \sum_{k=0}^{n} \frac{2n! \times n!^2}{k!^2 \cdot (n-k)!^2 \cdot n!^2}
\]

\[
= 1/4^{2n} \frac{2n!}{(n!)^2} \sum_{k=0}^{n} \frac{n!^2}{k!^2 \cdot (n-k)!^2}
\]

\[
= 1/4^{2n} \binom{2n}{n} \sum_{k=0}^{n} \binom{n}{k}^2
\]

But \( \sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n} \)

\[
\therefore \quad U_{2n} = 1/4^{2n} \binom{2n}{n}^2
\]  

(3.14)

In three dimension the particle moves along the three coordinates x, y and z in unit steps. There are 6 directions and so each step can have a probability 1/6. In symmetric random walk all the 6 directions have the same probability. A return to origin is possible only when the number of steps taken in the +ve x, y and z directions is equal to that taken in the -ve x, y and z directions, respectively.

Let \( k \) be the number of steps in the positive x-direction and \( j \) steps in the positive y-direction and \((n-j-k)\) steps in the positive z-direction. For a return to origin the same number of steps should be taken in the negative directions.
The summation extending over all $j, k$ with $(j + k) \leq n$

\[
U_{2n} = 6^{-2n} \sum_{j,k} \frac{2n}{j!^2 k!^2 \{(n-j-k)!\}^2}
\]  \hspace{1cm} (3.15)

\[
U_{2n} = \frac{1}{2^{2n}} \binom{2n}{n} \sum_{j,k} \left[ \frac{n!}{3^n j! k! (n-j-k)!} \right]^2
\]  \hspace{1cm} (3.16)

3.5 Probability as a measure

A random variable is a function: a function whose numerical values are determined by chance or a random variable is a real valued function on the measure space $X$. Halmos provided the justification for treating probability as a measure. Probability theory consists of the study of Boolean algebras of sets. We can translate the physical terminology into set theoretic terminology using the right notation.

An event is a set and its opposite event is the complementary set. Mutually exclusive events are disjoint sets. An event consisting of simultaneous occurrence of two other events is a set obtained by intersecting two other sets.

If $A$ is an event, the complementary event is denoted by $A'$. For example, in the gambling games of rolling an ordinary six-sided die and observing the number $x$ (= 1, 2, 3, 4, 5 or 6) showing on the top of the die, the number $x$ is an event. Event is one of the possible outcomes of some physical experiment. Let us consider a set $A$ where the event $x$ is an even number. Then $A = \{2, 4, 6\}$ and $A' = \{1, 3, 5\}$. To represent an event which occurs at least in one of the two events $A$ and $B$, we use $A \cup B$ and to represent an event which occurs in both the events $A$ and $B$ we use $A \cap B$. Thus if $A = \{2, 3, 5\}$ and $B = \{1, 2, 4\}$ then $A \cup B = \{1, 2, 3, 4, 5\}$ and $A \cap B = \{2\}$ and $A - B = \{3, 5\}$. 

Prepared by BeeHive Digital Concepts Cochin for Mahatma Gandhi University Kottayam
Another assumption in the theory of probability used here is that the system of events is closed under the formation of countably infinite unions, that the Boolean algebra is in fact a \( \sigma \)-algebra. Probability is a numerically valued function \( \mu \) of events \( E \), that is of sets of a Boolean \( \sigma \)-algebra. In other words numerical probability is a measure \( \mu \) on a Boolean \( \sigma \)-algebra \( S \) of subsets of a set \( X \), such that \( \mu(X) = 1 \).

3.5.1 Rings and algebras

To understand about rings and algebra we should have an idea about class. Class is a word used for a set of sets. If \( X \) is the Euclidean plane and \( E_y \) is the class of all intervals on the horizontal line at distance \( Y \) from the origin. Then each \( E_y \) is a class and the set of all these classes is a collection.

![Figure 3.5. Representation of a class of intervals in a plane.](image)

Each interval is a set and \( \bigcup E_y \) is the line itself, i.e., class of intervals.
(a) **Boolean ring (a ring of sets)**

It is a non empty class \( R \) of sets such that if \( E \in R \) and \( F \in R \) then \( E \cup F \in R \) and \( E - F \in R \).

In other words a ring is a non empty class of sets which is closed under the formation of unions and differences.

The empty set belongs to every ring \( R \) for if \( E \in R \), then

\[
0 = E - E \in R
\]

Another statement of ring is that it is a non empty class of sets closed under the formation of unions and symmetric differences.

(b) **Boolean algebra (algebra)**

It is a non empty class \( R \) of sets such that

(a) if \( E \in R \) and \( F \in R \), then \( E \cup F \in R \) and (b) if \( E \in R \), then \( E' \in R \).

Since \( E - F = E \cap F' = (E' \cup F)' \)

From the above identity it is concluded that every algebra is a ring. An algebra may be characterised as a ring containing \( X \). Since

\[
E' = X - E
\]

So it is clear that every such ring is an algebra.

Conversely, \( R \) is an algebra and \( E \in R \)

then, \( X = E \cup E' \in R \)

(c) **\( \sigma \)-ring**

\( \sigma \)-ring is a non empty class \( S \) of sets such that (a) if \( E \in S \) and \( F \in S \), then \( E - F \in S \) and (b) if \( E_i \in S \), \( i = 1, 2, \ldots \) then \( \bigcup_{i=1}^{\infty} E_i \in S \). Equivalently a \( \sigma \)-ring is a
ring closed under the formation of countable unions. If $S$ is a \( \sigma \)-ring and if \( E_i \in S \), \( i = 1, 2, \ldots \) and

\[
E = \bigcup_{i=1}^{\infty} E_i
\]

then the identity

\[
\bigcap_{i=1}^{\infty} E_i = E - \bigcup_{i=1}^{\infty} (E - E_i)
\]

shows that

\[
\bigcap_{i=1}^{\infty} E_i \in S,
\]

i.e., a \( \sigma \)-ring is closed under the formation of countable intersections.

### 3.5.2 Measure on rings

A set function is a function whose domain of definition is a class of sets. An extended real valued set function \( \mu \) defined on a class \( E \) of sets is additive if, whenever

\[
E \in E, F \in E, E \cup F \in E \text{ and } E \cap F = 0
\]

then

\[
\mu(E \cup F) = \mu E + \mu F
\]

An extended real valued set function \( \mu \) defined on a class \( E \) is finitely additive if, for every finite, disjoint class \( \{E_1, E_2, \ldots, E_n\} \) of sets in \( E \) whose union is also in \( E \), we have,

\[
\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n)
\]
A measure is an extended real valued, non-negative and countably additive set function, \( \mu \), defined on a ring \( R \), and such that \( \mu(0) = 0 \).

Since \( \bigcup_{i=1}^{n} E_i = E_1 \cup E_2 \cup \ldots \cup E_n \cup 0 \cup 0 \ldots \)

A measure is always finitely additive.

If \( \mu \) is a measure on a ring \( R \), a set \( E \) in \( R \) is said to have finite measure if \( \mu(E) < \infty \); the measure of \( E \) is \( \sigma \)-finite if there exists a sequence \( \{E_n\} \) of sets in \( R \) such that

\[
E = \bigcup_{n=1}^{\infty} E_n \quad \text{and} \quad \mu(E_n) < \infty, \quad n = 1, 2, \ldots
\]

**Totally finite measure:** If the measure of every set \( E \) in \( R \) is finite (or \( \sigma \)-finite), the measure \( \mu \) is called finite (or \( \sigma \)-finite) on \( R \). If \( R \) is an algebra, \( x \in R \) and if \( \mu(x) \) is finite or \( \sigma \)-finite, then \( \mu \) is called *totally finite* or *totally \( \sigma \)-finite*, respectively. The measure \( \mu \) is called complete if the conditions

\[
E \in R, \quad F \subseteq E \quad \text{and} \quad \mu(E) = 0 \quad \text{imply that} \quad F \in R.
\]

**3.5.3 A measurable space**

Measurable space is a set \( X \) and a \( \sigma \)-ring \( S \) of subsets of \( X \) with the property that \( \cup S = X \).

\( \mu \) defined on measurable space \( (X, S) \) \( \mu \) is a infinite valued measure—there exists certain elements in \( X \), whose images are \( +\infty \) or \( -\infty \) so that \( \mu^{-1}\{\infty\} \), \( \mu^{-1}\{-\infty\} \) are subsets of \( X \); if they are members of \( S \) then they are measurable sets.
Measure space

A measure space is a measurable space \((X, \mathcal{S})\) and a measure \(\mu\) on \(\mathcal{S}\) and it is usually written as \((X, \mathcal{S}, \mu)\). The measure space \(X\) is called totally finite, \(\sigma\)-finite or complete, according as the measure \(\mu\) is totally finite, \(\sigma\)-finite or complete.

3.5.4 Probability

Probability is a measure \(\mu\) on a Boolean \(\sigma\)-algebra \(\mathcal{S}\) of subsets of a set \(X\) such that \(\mu(X) = 1\).

A random variable is a measurable function on the measure space \((X, \mathcal{S}, \mu)\) for \(X\) where \(\mu\) is the probability measure.

3.6 The classification of random walk

The simplest definition of random walk is an analytical one. But in the case of problems concerning the transition function measure theory is used.\(^1\)__\(^6\)

3.6.1 Transition function

\(\mathbb{R}\) denote the space of \(d\)-dimensional integers, i.e., \(\mathbb{R}\) is the set of ordered \(d\)-tuples (lattice points). In the case of random walk \(\mathbb{R}\) is the state space of the random walk.

\[
x = (x^1, x^2, x^3, ... x^d), \quad x^i = \text{integer for } i = 1, 2, ... d \quad (3.17)
\]

For each pair \(x\) and \(y\) in \(\mathbb{R}\) we define a real number \(P(x, y)\) and the function \(P(x, y)\) will be called the transition function of the random walk. It is required to have the properties.
\[ 0 \leq P(x, y) = P(0, y-x), \quad \sum_{x \in \mathbb{R}} P(0, x) = 1 \] (3.18)

There is a spatial homogeneity for the above properties and is expressed by \( P(x, y) = P(0, y-x) \), where \((y-x)\) is the point in \( \mathbb{R} \) with coordinates \( y^i-x^i \), \( i = 1, 2, \ldots, d \). The transition function can be determined by a single function \( P(x) = P(0, x) \) on \( \mathbb{R} \) with the properties.

\[ 0 \leq P(x), \quad \sum_{x \in \mathbb{R}} P(x) = 1 \] (3.19)

Thus specifying a transition function is equivalent to specifying a probability measure on \( \mathbb{R} \).

A random walk is defined as a function \( P(x, y) \) possessing property,

\[ 0 \leq P(x, y) = P(0, y-x), \quad \sum_{x \in \mathbb{R}} P(0, x) = 1 \]

defined for all pairs \( x, y \) in a space of lattice points \( \mathbb{R} \). A random walk is said to be \( d \)-dimensional if the dimension of \( \mathbb{R} \) is \( d \).

For simple random walks, the Euclidean distance of the points \( x \) from the origin is given by

\[ |x| = \left[ \sum_{i=1}^{d} (x^i)^2 \right]^{1/2} \] (3.20)

where \( \mathbb{R} \) is \( d \)-dimensional. Then \( P(0, x) \) defines \( d \)-dimensional simple random walk if

\[ P(0, x) = \begin{cases} 1/2d & \text{when } |x| = 1 \\ 0 & \text{when } |x| \neq 1. \end{cases} \]
When \( d = 1 \), \( P(0,1) = p \), \( P(0,-1) = q \), \( p \geq 0 \), \( q \geq 0 \), \( p + q = 1 \), we shall call \( P(x,y) \) the transition function of Bernoulli random walk.

\( P(0,x) \) denote the probability of a “one step” transition from \( 0 \) to \( x \) and \( P_n(0,x) \) the probability of an “\( n \)-step” transition from \( 0 \) to \( x \). Suppose that \( n \) and \( x \) are both even or both odd and that, \( |x| \leq n \) (otherwise \( P_n(0,x) \) will be zero). Then \( P_n(0,x) \) should be the probability of \( \frac{1}{2}(x+n) \) successes in \( n \) Bernoulli trials, where the probability of success is \( p \) and that of failure is \( q \).

For an arbitrary random walk,

\[
P_n(0,x) = p^{(n+x)/2} q^{(n-x)/2} \binom{n}{(n+x)/2}
\]

(3.21)

when the sum \( (n+x) \) is even and \( |x| \leq n \), and \( P_n(0,x)=0 \).

\[
P_n(0,x) = \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \mathbb{R}} \ldots \sum_{x_{n-1} \in \mathbb{R}} P(0,x_1) P(x_1,x_2) \ldots P(x_{n-1},x) \quad (3.22)
\]

This result is derived from the following definition, i.e., for all \( x,y \) in \( \mathbb{R} \),

\[
P_0(x,y) = \delta(x,y) = 1 \text{ if } x=y,
\]

\[
= 0 \text{ otherwise}
\]

and \( P_1(x,y) = P(x,y) \).

\[
P_n(x,y) = \sum_{x_i \in \mathbb{R}, i=1 \ldots n-1} P(x,x_1) P(x_1,x_2) \ldots P(x_{n-1},y) \quad n \geq 2
\]

(3.23)

the sum extends as briefly indicated, over all \( (n-1) \) tuples \( x_1 \ x_2 \ldots x_{n-1} \) of points in \( \mathbb{R} \). From this definition the following results are derived.

For all \( x,y \) in \( \mathbb{R} \),

\[
P_{n+m}(x,y) = \sum_{t \in \mathbb{R}} P_n(x,t) \ P_m(t,y) \text{ for } n \geq 0, \ m \geq 0
\]

(3.24)
\[ \sum_{y \in \mathbb{R}} P_n(x, y) = 1, \ P_n(x, y) = P_n(0, y-x) \] for \( n \geq 0 \) \hspace{1cm} (3.25)

The probability \( P_n(x, y) \) represents the probability that a 'particle' executing the random walk and starting at the point \( x \) at time 0 will be at the point \( y \) at time \( n \). The function \( F_n(x, y) \) is not a transition function. It is the probability, of the first visit to the point \( y \) at the time \( n \), starting from the point \( x \). The following definitions are concerning the function \( F_n(x, y) \).

\[ F_0(x, y) = 0, \ F_1(x, y) = P(x, y) \] \hspace{1cm} (3.26)

\[ F_n(x, y) = \sum_{x_i \in \mathbb{R} \setminus \{y\}} P(x, x_1) P(x_1, x_2) \ldots P(x_{n-1}, y) \] \hspace{1cm} where \( n \geq 2 \) \hspace{1cm} (3.27)

for all \( x, y \) in \( \mathbb{R} \).

\( \{y\} \) - subset of \( \mathbb{R} \) consisting the element \( y \) and \( \mathbb{R} \setminus \{y\} \) denote the state space \( \mathbb{R} \) with the point \( y \) excluded.

There are 3 important properties concerning \( F_n(x, y) \). They are

1. \[ F_n(x, y) = F_n[0,(y-x)] \] \hspace{1cm} (3.28)

2. \[ \sum_{k=1}^{n} F_k(x, y) \leq 1 \] \hspace{1cm} (3.29)

3. \[ P_n(x, y) = \sum_{k=1}^{n} F_k(x, y) P_{n-k}(y, y) \] \hspace{1cm} (3.30)

for \( n \geq 1 \) and \( x, y \in \mathbb{R} \).

The next function to be defined is \( G_n(x, y) \). This corresponds to the expected number of visits of random walk, starting from \( x \) and visits the point \( y \) within the time \( n \). This function is actually the expectation, being defined as the sum of the probabilities, i.e.,

\[ G_n(x, y) = \sum_{k=0}^{n} P_k(x, y), \ n = 0, 1, \ldots, \ x, y \in \mathbb{R} \] \hspace{1cm} (3.31)
From this it is proved that

\[ G_n(x,y) \leq G_n(0,0) \text{ for } n \geq 0 \text{ and all } x,y \in \mathbb{R} \quad (3.32) \]

3.6.2 Recurrent and transient random walks

Random walks can be classified into recurrent and transient (i.e., non recurrent). According to the theory of Markov chains, the probability of a return to the starting point is one for some but not all points of the state space. Such points are then called recurrent or persistent. The points with return probability less than one are transient.

The sum \( \sum_{k=1}^{n} F_k(0,0) \) represents the probability of a return to the starting point before or at the time \( n \). The sequence of the sums \( \sum_{k=1}^{n} F_k(0,0) \) is non decreasing as \( n \) increases, and they are bounded by one according to equation

\[ \sum_{k=1}^{n} F_k(x,y) \leq 1. \]

This function has a limit as \( n \to \infty \) and at the limiting condition it is represented as \( \mathbf{F} \) and \( \mathbf{F} \leq 1 \). The random walk is recurrent if \( \mathbf{F} = 1 \) and transient if \( \mathbf{F} < 1 \).

The limit of the monotone sequence \( G_n(0,0) \) as \( n \to \infty \) is \( \mathbf{G} \). \( \mathbf{G} \) may be finite or infinite.

\[ G(x,y) = \sum_{n=0}^{\infty} P_n(x,y) \leq \infty; \quad F(x,y) = \sum_{n=1}^{\infty} F_n(x,y) \leq 1 \quad (3.33) \]

\[ G_n(0,0) = G_n; \quad G(0,0) = \mathbf{G}; \quad F_n(0,0) = F_n; \quad F(0,0) = \mathbf{F} \quad (3.34) \]

Another result derived from the above relations is Theorem 1.
Theorem 1

\[ G = \frac{1}{1-F} \] with the condition that \( G = +\infty \) when \( F=1 \) and \( F=1 \) when \( G = +\infty \).

Proof

\[ G_n(0,0) = \sum_{k=1}^{n} P_k(0,0) \]

We know that

\[ P_n(x,y) = \sum_{k=1}^{n} F_k(x,y) P_{n-k}(y,y) \]

Putting \( x=y=\text{origin} \), we can write

\[ P_n(0,0) = \sum_{k=0}^{n} F_k(0,0) P_{n-k}(0,0) \] (a)

But \( F_k(0,0) = F_k \) by the convention used here

\[ \therefore \quad P_n(0,0) = \sum_{k=0}^{n} F_k P_{n-k}(0,0) \quad n \geq 1 \]

\[ \sum_{n=1}^{n} P_n(0,0) = \sum_{n=1}^{m} \sum_{k=0}^{n} F_k P_{n-k}(0,0) \]

\[ n = 1, 2, \ldots m \]

Adding \( P_0(0,0) \) on both sides of the above equation.

\[ P_0(0,0) + \sum_{n=1}^{m} P_n(0,0) = \sum_{k=0}^{n} F_k \sum_{n=1}^{m} P_{n-k}(0,0) + P_0(0,0) \]

But \( P_0(0,0) = \sum_{n=1}^{m} P_n(0,0) = G_m \) and \( P_0(0,0) = 1 \)
The above equation becomes

\[ G_m = \sum_{k=0}^{n} F_k \sum_{n=1}^{m} P_{n-k}(0,0) + 1 \]

\[ G_m = \sum_{k=0}^{m} F_k G_{m-k} + 1 \tag{b} \]

Since \( n = 1, \ldots, m \) and \( m \geq 1 \). In the limiting case \( m \to \infty \).

\[ G_m = G \]

\[ G = 1 + \lim_{m \to \infty} \sum_{k=0}^{m} F_k G_{m-k} \geq 1 + G \sum_{k=0}^{N} F_k \]

Since \( N \) is a finite number which will always be less than \( \infty \), we can write

\[ G \geq 1 + GF \]

This proves that \( G = +\infty \) when \( F = 1 \). Since the inequality \( G \geq 1 + G \) has no finite solutions.

From equation (b) we get

\[ 1 = G_m - \sum_{k=0}^{m} G_k F_{m-k} \geq G_m - G_m \sum_{k=0}^{m} F_{m-k} \geq G_m(1-F) \]

So that \( 1 \geq G(1-F) \) which shows that \( G < \infty \) when \( F < 1 \).

This proves that \( G(1-F) = 1 \) or

\[ G = \frac{1}{(1-F)} \tag{3.35} \]

**Theorem 2**

For every random walk,

\[ \lim_{n \to \infty} \frac{G_n(x,y)}{G_n(0,0)} = F(x,y) \] wherever \( x \neq y \) \tag{3.36}
Proof

Consider the special case when $y = 0$

$$G_n(x,0) = \sum_{k=0}^{n} P_k(x,0) \quad n = 0, 1, \ldots, x, y \in \mathbb{R}$$

$$G_n(x,0) = P_0(x,0) + \sum_{k=1}^{n} P_k(x,0)$$

Since $P_0(x,y) = \delta(x,y)$

$\delta(x,y)$ is the kronecker delta which has a property

$\delta(x,y) = 0$ when $x \neq y$

$\delta(x,y) = 1$ when $x = y$

We get $P_0(x,0) = \delta(x,0)$

$$\therefore G_n(x,0) = \delta(x,0) + \sum_{k=1}^{n} P_k(x,0)$$

Using Equation 3.30 which states that

$$P_n(x,y) = \sum_{k=1}^{n} F_k(x,y) P_{n,k}(y,y)$$

$G_n(x,0)$ can be written as

$$G_n(x,0) = \delta(x,0) + \sum_{k=1}^{n} \sum_{j=1}^{k} F_j(x,y) P_{kj}(y,y)$$

Here $y = 0$.

$$\therefore G_n(x,0) = \delta(x,0) + \sum_{k=1}^{n} \sum_{j=1}^{k} F_j(x,0) P_{kj}(0,0)$$

A simple interchange of the order of summation

put $i = k-j$, $0 \leq j \leq k \leq n$

$0 \leq k-j \leq n-j \quad j \leq k$
or \( 0 \leq i \leq n \) \( j \leq n \) 
\( k \leq n \) 
\( i \leq n - j \)

\[ G_n(x,0) = \delta(x,0) + \sum_{j=0}^{n} \sum_{i=0}^{n-j} F_{i}(x,0) P_{j}(0,0) \]
\[ = \delta(x,0) + \sum_{j=0}^{n} P_{j}(0,0) \sum_{i=0}^{n-j} F_{i}(x,0) \]

\[ G_n(0,0) = \sum_{j=1}^{n} P_{j}(0,0) \]

\[ \frac{G_n(x,0)}{G_n(0,0)} = \frac{\delta(x,0)}{\sum_{j=0}^{n} P_{j}(0,0)} + \frac{\sum_{i=0}^{n-j} F_{i}(x,0)}{\sum_{j=0}^{n} P_{j}(0,0)} \]

when \( x \neq y \), \( \delta(x,0) = 0 \)

In the transient case \( \sum_{j=0}^{n} P_{j}(0,0) \) has a finite limit and so

\[ \lim_{n \to \infty} \frac{G_n(x,0)}{G_n(0,0)} = \lim_{n \to \infty} \sum_{i=0}^{n-j} F_{i}(x,0) \]

But \[ \lim_{n \to \infty} \sum_{k=0}^{n} F_{k}(x,0) = F(x,0) \]

\[ \therefore \lim_{n \to \infty} \frac{G_n(x,0)}{G_n(0,0)} = F(x,0) \]

In the recurrent case

\[ \sum_{j=0}^{n} P_{j}(0,0) \to \infty \]

So we cannot cancel the numerator and the denominator terms.