NATIONAL INSTITUTE OF TECHNOLOGY, SILCHAR

Department of Mathematics

Project Report on,

"RANDOM WALK AND PROPERTIES ASSOCIATED WITH IT SUCH AS MEAN, VARIANCE, AUTO-CORRELATION FUNCTION, AUTO-COVARIANCE FUNCTION AND ITS ONE STEP TRANSITION PROBABILITIES. PERFORMING CHECKS FOR MARKOV CHAIN. DISCUSSION OF 'GAMBLER'S RUIN PROBLEM' WITH EXAMPLES"

Subject : Mathematics-IV (MA-202)

Submitted by,
Name : Arvinder Singh
Scholar ID : 18-1-5-126
B.Tech 4th SEM (CSE)

July 8, 2020

ABSTRACT

In this project, I have given a concise overview of the mathematical analysis of the concept of 'simple random walk'. The various properties of this random or stochastic process is discussed in full detail.

Graphical visualization is given for better understanding and intuition. Definitions of various subject matters are given as and when required with proper citing of the equations that are used throughout the project.

This project was developed using open source software LaTeX.

Contents

1 Introduction

A *random walk* is a discrete time *stochastic* or *random* process, with discrete time random variable, which can be denoted as the following:

$$X_0, X_1, X_2, \dots \tag{1}$$

A *random process* is a time-varying function that assigns the outcome of a random experiment to each time instant X_t.The above is also sometimes called a 1-*dimensional* simple random variable.
Let us now define a independent identically distributed random variable Y_i as follows:

$$Y_i = \begin{cases} 1, & \text{with probability } \frac{1}{2} \\ -1, & \text{with probability } \frac{1}{2} \end{cases}$$

If $P(Y_i)$ denotes the probability of Y_i taking a particular value then it can also be defined as the following:

$$P(Y_i = 1) = \frac{1}{2}, P(Y_i = -1) = \frac{1}{2} \tag{2}$$

As I have already mentioned that *random walk* is a *discrete time-parameter* stochastic process, with discrete random variables taking values by the following equation,

$$X_t = \sum_{i=1}^{t} Y_i \ \ , \ \ X_0 = 0 \tag{3}$$

We can plot the above equation with t taken in $X-axis$ and X_t taken in $Y-axis$.

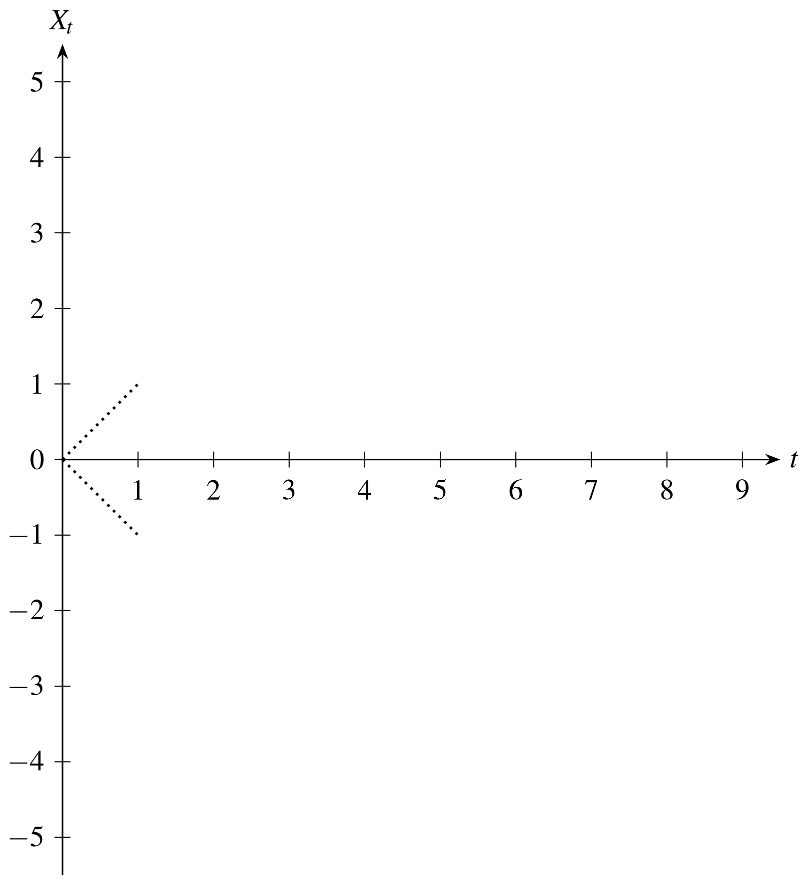

Figure 1: This figure shows all probable directions in which the event X_1 can turn out to be.

Suppose the event in the '**Figure 1**' turns out as follows:

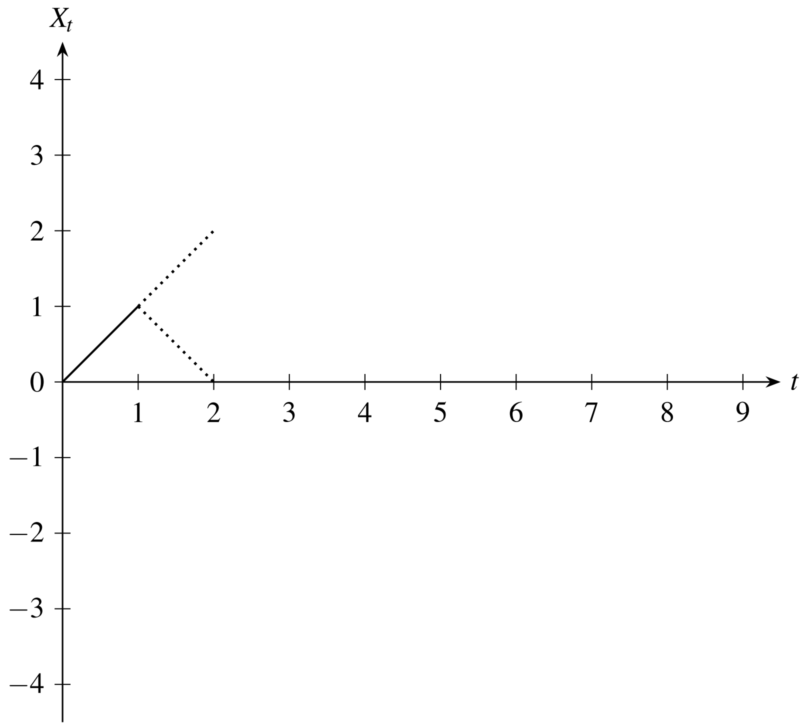

Figure 2: This figure shows all probable directions in which the event X_2 can turn out, we can observe that it solely depends upon X_1.

This gives us enough intuition to develop an example for random walk. One such example is given below:

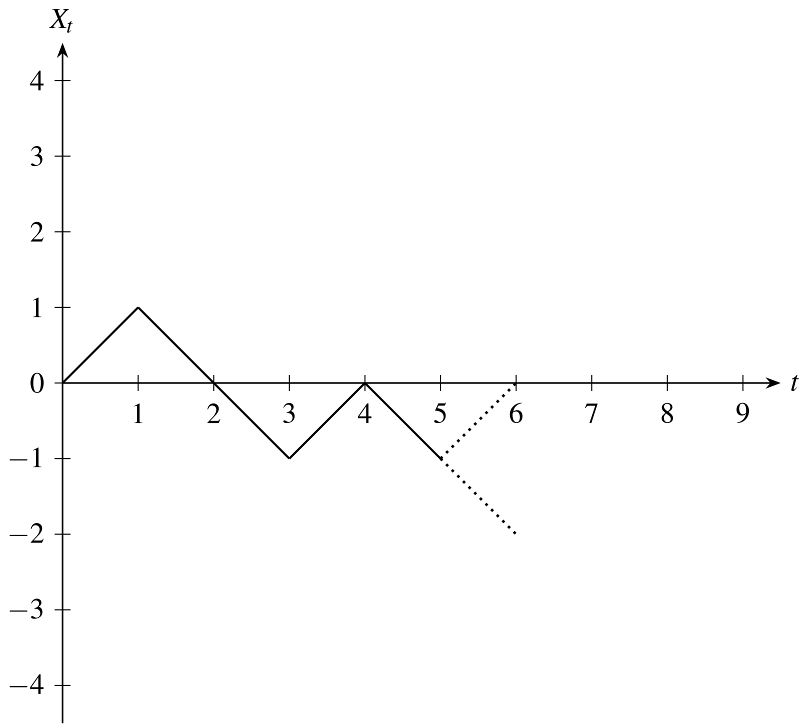

Figure 3: This figure shows one of many possible outcomes.

We have seen visually and understood the intuition of how this process develops. We now try to understand why this process is called *'Random Walk'*.

If we take a look at the projection of the **Figure 3**, about $X-axis$ it is not interesting as it only progresses in one direction, on the other hand the projection of it about $Y-axis$ is interesting. It basically oscillates randomly back and forth. It is like someone walking randomly on a one dimensional path changing directions. Hence, it is named so.

2 Theory and Properties

The event Y_i as represented by equation (2) can be an event such as : Picking a *red* or a *blue* ball from a covered box containing *two* ball among which one is *red* and the other is *blue*, rolling a *dice* and the outcome to be an *even* or a *odd* number, tossing a coin and outcome to be either *heads* or *tails*,etc.
Here we demonstrate with the help of a table how an example graph such as the one already shown in the **Figure 3** can be constructed. Here, we consider the first example (*red and blue ball*) of the event Y_i *(again, it is independent identically distributed)* and generalize the probability as p and q as shown below. We have, $p+q=1$

$$Y_i = \begin{cases} 1, & \text{if the ball drawn is blue(B), probability=}p \\ -1, & \text{if the ball drawn is red(R), probability=}q \end{cases} \tag{4}$$

t	0	1	2	3	4	5	6	7	8	9
Draw		R	R	B	R	B	B	R	B	B
Y_t	0	-1	-1	1	-1	1	1	-1	1	1
X_t	0	-1	-2	-1	-2	-1	0	-1	0	1

Table 1: Table depicting each draw Y_t and state X_t

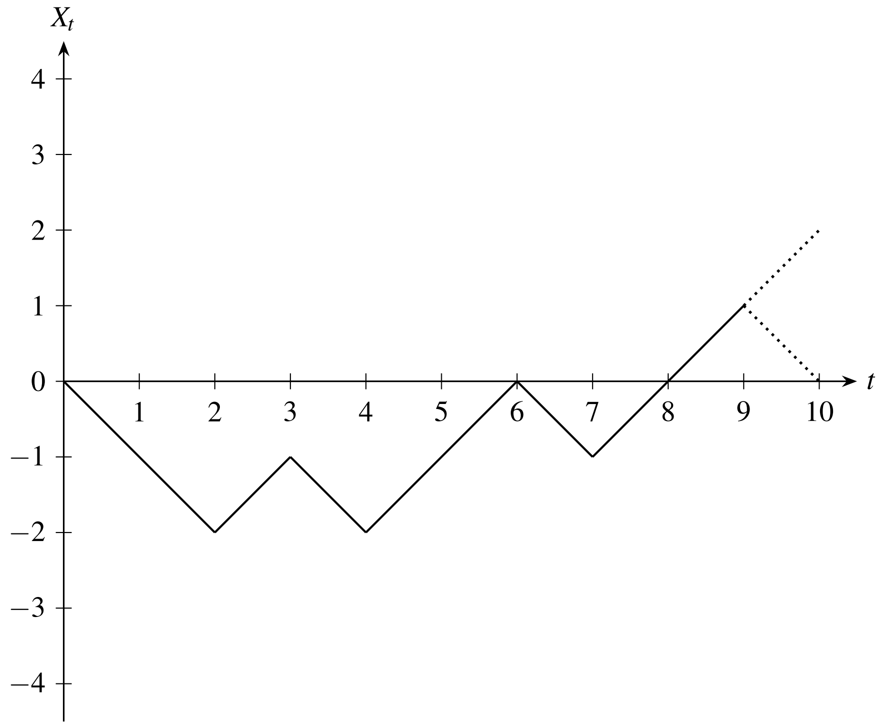

Figure 4: This figure depicts the values taken by X_t for set of outcomes as shown in the table.

We can now conclude that the sate space of random variable X_t is discrete and take the values belonging to the following set:

$$S = \{..., -3, -2, -1, 0, 1, 2, 3, ...\} \tag{5}$$

We also develop a recursive function for the random process X_t for an event as an example *equation* (4).
Using *equation* 3 we have,

$$X_t = Y_t + \sum_{i=1}^{t-1} Y_i \tag{6}$$

$$\implies X_t = Y_t + X_{t-1}, (From\ \ 3) \tag{7}$$

equation 7 is defined for $t \geq 1$. Also, $X_0 = 0$ as already defined.
Also, here this *simple random walk* is said to be unrestricted, as there is no boundary condition on the values possible for X_t.

2.1 Probability Distribution

Here we basically try to develop a model that can give us probability for different values of X_t. We use the same notation as described in *equation* (2), except that we generalize it for value i. In other words we find the value of,

$$P(X_t = i) \tag{8}$$

The above can be read as, "Probability that $X_t = i$ at time t".
Definitely,

$$\begin{gathered} X_t \leq t \\ \implies X_t = i \leq t \\ \implies i \leq t \end{gathered}$$

The above also implies that,

$$P(X_t = i > t) = 0 \tag{9}$$

Also, setting initial condition of the *random walk* gives,

$$P(X_t = 0) = 1 \tag{10}$$

We introduce some more notations, let α_t denote the total number of $+1$ and β_t denote the total number of -1 for a particular X_t at time t,
Obviously the following conditions holds,

$$\alpha_t + \beta_t = t \tag{11}$$

$$\alpha_t - \beta_t = X_t = i \tag{12}$$

Solving the *equations* (11) and (12) we have,

$$\alpha_t = \frac{1}{2}(t+i), \beta_t = \frac{1}{2}(t-i) \tag{13}$$

Now, we find the probability given by *equation* (8). We ask ourselves a question, "How many ways are there for X_t to be i?". The answer is the number of ways in which we can select the number of α_t, for that particular i, the relation of which is given by *equation* (13). Then we multiply the probabilities of having that number of 1s as in binomial distribution.

$$\boxed{P(X_t = i) = {}^nC_{\frac{1}{2}(t+i)} p^{\frac{1}{2}(t+i)} q^{\frac{1}{2}(t-i)} = {}^nC_{\alpha_t} p^{\alpha_t} q^{\beta_t}} \tag{14}$$

2.2 Mean

Before we find mean of X_t, we need to find the mean of Y_i, we know that it is independent identically distributed random variable as mentioned before,
The formula of mean is given by,

$$\mu_Y = E(Y) = \Sigma x_k p_Y(x_k) \tag{15}$$

where, in above formula $p_Y(x_k)$ is probability of $Y = x_k$.
The *mean*, or *expectation* of Y_i is given by,

$$E(Y_i) = \Sigma x_k p_{Y_i}(x_k) \tag{16}$$

The probabilities and corresponding values are given by *equation* (4),

$$E(Y_i) = (1)(p) + (-1)(q) = p - q \tag{17}$$

Now, we have sufficient set-up to derive the mean of X_t,

$$E(X_t) = E(\sum_{i=1}^{t} Y_i) \tag{18}$$

Since all the Y_i's are independent and identical, we can write the above equation as,

$$E(X_t) = tE(Y_i) \tag{19}$$

From *equation* (17),

$$\boxed{E(X_t) = t(p-q)} \tag{20}$$

2.3 Variance

Before we find variance of X_t, we need to find the variance of Y_i,
Before we calculate the variance of Y_i we need to calculate $E(Y^2)$,

$$E(Y^2) = \Sigma x_k^2 p_Y(x_k) \tag{21}$$

where, in above formula $p_Y(x_k)$ is probability of $Y = x_k$.

$$\implies E(Y_i^2) = \Sigma x_k^2 p_{Y_i}(x_k) \tag{22}$$

The probabilities and corresponding values are given by *equation* (4),

$$\implies E(Y_i^2) = (1)^2(p) + (-1)^2(q) = p + q = 1 \tag{23}$$

Now, we have sufficient set-up to derive the variance of Y_i,

$$Var(Y_i) = E(Y_i^2) - [E(Y_i)]^2 \tag{24}$$

Putting the values of $E(Y_i^2)$ and $E(Y_i)$ from *equation* (23) and (17) we have,

$$\implies Var(Y_i) = 1 - (p-q)^2$$
$$\implies Var(Y_i) = 1 - [(p+q)^2 - 4pq]$$
$$\implies Var(Y_i) = 1 - [1 - 4pq]$$
$$\implies Var(Y_i) = 4pq$$

$$\boxed{Var(Y_i) = 4pq} \tag{25}$$

We can now come up with variance with X_t, given by,

$$Var(X_t) = Var(\sum_{i=1}^{t} Y_i) \tag{26}$$

Using the property of variance that $Var(A+B) = Var(A) + Var(B)$, also keeping in mind the fact that Y_i is independent identically distributed random variable we have,

$$Var(X_t) = tVar(Y_i) \tag{27}$$

Using *equation* (25) we have,

$$\boxed{Var(X_t) = 4tpq} \tag{28}$$

2.4 Auto-correlation function

The auto-correlation function for the two parameters t and s is given by,

$$R_X(t,s) = E[X_t X_s]$$

$$\implies R_X(t,s) = E[\sum_{i=1}^{t} Y_i \sum_{j=1}^{s} Y_j] \tag{29}$$

Using properties of expectation we have,

$$\implies R_X(t,s) = \sum_{i=1}^{t}\sum_{j=1}^{s} E[Y_i Y_j] \tag{30}$$

Using some basic rules and principle it can be shown that the above is equal to,

$$\implies R_X(t,s) = \sum_{i=1}^{min(t,s)} E[Y_i^2] + \sum_{i=1}^{t}\sum_{\substack{j=1 \\ i\neq j}}^{s} E[Y_i][Y_j] \tag{31}$$

Using *equations* (17) and (23) and using *inclusion-exclusion* principle we have,

$$\implies \boxed{R_X(t,s) = min(t,s) + [ts - min(t,s)](p-q)^2} \tag{32}$$

For a special case when, $p = q = \frac{1}{2}$ we have,

$$R_X(t,s) = min(t,s)$$

2.5 Auto-covariance function

The auto-covariance function for the two parameters t and s is given by,

$$K_X(t,s) = R_X(t,s) - \mu_X(t)\mu_X(s) \tag{33}$$

Using *equations* (20) and (32) we have,

$$K_X(t,s) = min(t,s) + [ts - min(t,s)](p-q)^2 - ts(p-q)^2 \tag{34}$$

Solving the above equation we have,

$$\boxed{K_X(t,s) = min(t,s)[1-(p-q)^2]} \tag{35}$$

3 Proving Random Walk is a Markov Chain

From *equation* (3) we have,

$$X_t = \sum_{i=1}^{t} Y_i \ \ , \ \ X_0 = 0 \ \ , \ \ t \geq 1$$

Moreover t takes discrete values,

$$t \in \{1,2,3,...\}$$

By definition a discrete time parameter *Markov Chain* is characterized by the following equation,

$$P(X_{t+1} = i_{t+1}|X_t = i_t, X_{t-1} = i_{t-1}, ..., X_0 = i_0) = P(X_{t+1} = i_{t+1}|X_t = i_t) \tag{36}$$

The above equation clearly indicates that in a Markov chain the future state depends solely on the current state, irrespective of all the past state. It will not be accurate to say that it does not depend on the past, but it will be more theoretically correct to say that all the past information is characterized or condensed in current state. To prove that *Random Walk* X_t is indeed a Markov Chain we use a recursive function we already defined in *equation* (7),

$$X_t = Y_t + X_{t-1}$$

From above equation we can clearly see that it only summation of two terms, Y_t is independent random variable, and X_{t-1}. This, shows that X_t only depends only upon the previous value *i.e.,* X_{t-1}. This completes our proof.

4 One-step transition probability

The one step transition probability by definition is given by,

$$p_{ji} = P(X_t = i|X_{t-1} = j) = \begin{cases} p, & i = j+1 \\ q, & i = j-1 \\ 0, & \text{otherwise} \end{cases} \tag{37}$$

Moreover, $p + q = 1$,

The above equations means that if the previous state was j then the next state, depending upon probability could be $j+1$ or $j-1$.

Also, if we are to have a transition probability matrix the state space has to be finite.

5 Applications

The following are some of the applications of *random walk* in practical scenarios:

- The concept of *simple random walk* is used in *'physics'* where, the physical movement of particles as in *'Brownian motion'* is modeled.Also, of *molecules* of liquids and gases.
- In subjects such as *'economics'*, the *random walk* is used to model share prices and other factors.
- In Computer Science, *random walks* are used to estimate the size of World Wide Web.

These are applications only to mention a few. Next we will be discussing the *Gambler's Ruin Problem*.

6 Gambler's Ruin Problem

Let us assume that there are two gamblers, G_1 and G_2, also, let the gambler G_1 has m_1 and G_2 has m_2 rupees.

The rules of games go like this:"G_1 wins 1 rupees and G_2 loses 1 rupees with probability p and G_1 loses 1 rupees and G_2 wins 1 rupees with probability q. They play until one of them goes broke!"

The total capital money combining the two players is $M = m_1 + m_2$.

Also, let Y_i be independent identically distributed random variables, we have,

$$P(Y_i = 1) = p \tag{38}$$

$$P(Y_i = -1) = q \tag{39}$$

Also, $p + q = 1$ and $i \geq 1$

Let X_t be the amount of money the gambler G_1 has at time t. This quantity can be represented as:

$$X_t = \sum_{i=1}^{t} Y_i \quad , \quad X_0 = 0 \tag{40}$$

The above equation can be recursively defined as follows:

$$X_t = Y_t + X_{t-1}$$

The above is a similar recursive equation as given by *equation* (7), except that here we have *terminating* or *absorbing* states $X_t = M$ or $X_t = 0$, which means that the game terminates when any one of these conditions gets satisfied.

In the previous section we have already discussed the criteria for a random process to be a Markov Chain. So, form the **Section** 3, we infer that this process is Markov chain with a discrete *finite* state space contrary to the one defined by *equation* (5). The finite state space can be represented as follows,

$$S = \{0, 1, 2, ..., M\} \tag{41}$$

Additionally, this process has has absorbing states.

This Markov Chain X_t is also known as *simple random walk with absorbing states*.

We define the transition probabilities with respect to the conditions mentioned above,

$$\begin{gathered} p_{i,i+1} = P(X_t = i+1|X_{t-1} = i) = p \\ p_{i,i-1} = P(X_t = i-1|X_{t-1} = i) = q \\ p_{i,i} = P(X_t = i|X_{t-1} = i) = 0 \\ p_{0,0} = P(X_t = 0|X_{t-1} = 0) = 1 \\ p_{M,M} = P(X_t = M|X_{t-1} = M) = 1 \end{gathered}$$

Since the given state space is finite we can draw a *generalized transition probability matrix*.

A general matrix can be defined as follows:

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ M-1 \\ M \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & 2 & 3 & \dots & M-2 & M-1 & M \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} \end{array}$$

The matrix P defined as above is the transition probability matrix with the vertical marking row-wise indicates the initial state and the horizontal markings at the top of the matrix indicates the next state corresponding to the current state.

The matrix is solely filled with the transitions already shown.

6.1 Example

As an example we can think that,"Two friends Tom and Jim play a game. Tom has initially 20 rupees and Jim has 15. They play a game. They start tossing a coin. If heads appear Tom wins 1 rupees else he loses 1 and Jim wins 1 and vice-versa. This game continues until Tom has all the money or say capital money of 35 rupees or goes broke for 0 rupees!"

Also, it is worth noting that in *Casinos*, the probability of players winning a game is biased. This, is because the *Casino* also needs to make profit and if it would have been unbiased then they would not make any profit at all. So, finite mathematics comes into play!

6.2 Probability of Ruin and Win

Taking a general example for Gambling, let us denote $P(X_0 = i)$ be the probability that gambler say G_1 (*as in the general example given in the start of this section*) goes broke, *i.e.,* loses all his money when initially he has i rupees. Also, it has finite discrete state space like the one given by *equation* (41).

We can also say that, since this kind of random process has two absorbing state, namely $X_t = 0$ and $X_t = M$, the defined $P(X_0 = i)$ is the probability for absorption at state $X_t = 0$.

Also, let the probability of transition be as defined by *equation* (38) and (39) *i.e.,* gaining 1 rupees has probability of p and losing has probability of q. We can now define a recursive function for the same. Now for $i \in (0, M)$ we have the following recurrence relation,

$$P(X_0 = i) = pP(X_1 = i+1) + qP(X_1 = i-1) \tag{42}$$

For simplicity we can write the above for general time t as,

$$P(i) = pP(i+1) + qP(i-1) \tag{43}$$

The states above can be read as,for the first term the "Gambler G_1 wins first round and subsequently in future loses all the money he had initially", the second term "Gambler G_1 loses first round and subsequently in future loses all the money he had initially".
Rearranging the equation we have,

$$P(i+1) - \frac{1}{p}P(i) + \frac{q}{p}P(i-1) = 0 \ \ , \ \ i \in (0, M) \tag{44}$$

By the construct of the problem we have, $P(0) = 1$ and $P(M) = 0$. *equation* is a homogeneous linear constant coefficient difference equation, using the following identity that $P(i+1) = \lambda P(i)$ and recursively implementing it we get,

$$\lambda^{i+1} - \frac{1}{p}\lambda^{i} + \frac{q}{p}\lambda^{i-1} = 0 \ \ , \ \ p+q = 1 \tag{45}$$

$$\lambda^2 - \frac{1}{p}\lambda + \frac{q}{p} = (\lambda - 1)(\lambda - \frac{q}{p}) = 0 \tag{46}$$

The above equation gives us the following solutions: $\lambda_1 = 1$ and $\lambda_2 = \frac{q}{p}$. The general solution of this difference equation is given by,

$$P(i) = k_1\lambda_1^i + k_2\lambda_2^i \ \ , \ \ p \neq q \tag{47}$$

Plugging the values of λ_1 and λ_2 in the above equation we have,

$$P(i) = k_1 + k_2(\frac{q}{p})^i \ \ , \ \ p \neq q \tag{48}$$

In the above equation k_1 and k_2 are arbitrary constants. Using the absorbing conditions we have,

$$k_1 + k_2 = 1$$
$$k_1 + k_2(\frac{q}{p})^M = 0$$

Solving the above two equations we have the values of k_1 and k_2 as,

$$k_1 = \frac{-(\frac{q}{p})^M}{1-(\frac{q}{p})^M}$$
$$k_2 = \frac{1}{1-(\frac{q}{p})^M}$$

Plugging the values of k_1 and k_2 in *equation* we have,

$$\boxed{P(i) = \frac{(\frac{q}{p})^i - (\frac{q}{p})^M}{1 - (\frac{q}{p})^M} \quad , \quad p \neq q} \tag{49}$$

For very large values of M we can assume that $M \to \infty$ then, based on whether $p < q$ or $p > q$, we have,

$$P(i) = \begin{cases} 1, & p < q \\ (\frac{q}{p})^i, & p > q \end{cases} \tag{50}$$

For the special case when $p = q$, let $\lambda = \frac{q}{p}$, then *equation*, (49) turns out to be,

$$P(i) = \frac{\lambda^i - \lambda^M}{1 - \lambda^M}$$

So, in the above $\lambda \to 1$. We can see that the equation above is of the form $\frac{0}{0}$. Hence, we can apply L'Hôpital's rule to the above we have,

$$P(i) = \frac{i\lambda^{i-1} - M\lambda^{M-1}}{-M\lambda^{M-1}}$$

Plugging in the value of $\lambda = 1$ in above we have,

$$P(i) = \frac{i - M}{-M}$$

$$\boxed{P(i) = 1 - \frac{i}{M}} \tag{51}$$

Therefore, above is the answer for the special case when $p = q = \frac{1}{2}$.
We now to find the probability of gambler G_1 winning if he has i rupees initially the answer is,

$$P'(i) = 1 - P(i) = \boxed{P(i) = \frac{1 - (\frac{q}{p})^i}{1 - (\frac{q}{p})^M}} \tag{52}$$

For very large values of M we can assume that $M \to \infty$ then, based on whether $p < q$ or $p > q$, we have,

$$P(i) = \begin{cases} 0, & p < q \\ 1 - (\frac{q}{p})^i, & p > q \end{cases} \tag{53}$$

For special case of $p = q = \frac{1}{2}$ we have,

$$\boxed{P(i) = \frac{i}{M}} \tag{54}$$

To summarize, the probability of gambler G_1 with i amount of money initially and losing all the money is given by *equation* (49) together with special cases as described by *equations* (50) and (51), for winning the corresponding equations are given by *equation* (52), (53), (54).

7 Conclusion

Starting with the introduction we have discussed the random process and how the random walk develops. We have seen the graphical construct of the *random walk*, taking some examples.

We saw the probability distribution of the random walk along with mean, variance, auto-correlation function, auto-covariance function and also seen the one-step transition probabilities.

We tried to understand why random walk is a Markov Chain, we have seen the recursive function and the dependency and relation among states and why it turns out to be Markov.

We have seen some applications of random process in particular *random walk* and in the end we have seen the 'Gambler's ruin problem' very concisely described with comprehensive mathematical analysis and discussed with some examples. We have also seen the probability of winning and losing with derivations of the corresponding equations.

References

[1] https://en.wikipedia.org/wiki/Random_walk

[2] https://en.wikipedia.org/wiki/Stochastic_process

[3] Random Variables and Random Processes, Schaum's Outline

[4] http://www.columbia.edu/~ks20/FE-Notes/4700-07-Notes-GR.pdf

[5] Stochastic Process-1,(MIT) https://www.youtube.com/watch?v=TuTmC8aOQJE&t=3322s

[6] Gambler's Ruin, Brendon Foltz https://www.youtube.com/watch?v=afIhgiHVnj0