Statistics and Probabilities In Markov Chain

Professor : *Dr. Mohammad A. Maddah-Ali*

Subject : *Intorduction to Markov Chains*

Author : *Ali Jafari*

Student ID : *97101434*

Spring 1399 - 1400

Sharif University
Of
Technology

Contents

1 Solutions :

1.1 Question 1 : Examples of Markov Chain

1. **Weather model :**

   The probabilities of weather conditions (modeled as either rainy or sunny), given the weather on the preceding day, , can be represented by a graph .

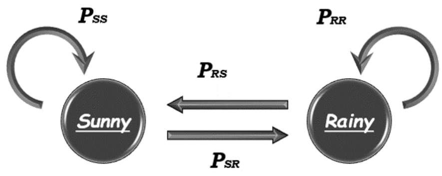

Figure 1: Weather State diagram

2. **Gamblers ruin :**

   Consider a tennis game in Which the the score has reached deuce . Suppose that computer wins with the probability P .

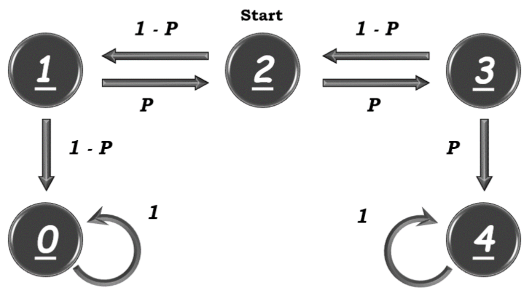

Figure 2: State diagram

3. **Random Walk :**

   A boy walks along a four-block stretch of a Road. If he is at corner 1, 2, or 3, then he walks to the left or right with equal probability. He continues until he reaches corner 4, which is a Park, or corner 0, which is his home. If he reaches either home or the park, he stays there.

   The transition matrix is like below :

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

1.2 Question 2

Show that :

$$\sum_{j=1}^{M} p_{ij} = 1 \qquad\qquad (\forall\, i \in \chi)$$

Proof :

$$\sum_{j=1}^{M} p_{ij} = \sum_{j=1}^{M} P[X_{n+1} = j | X_n = i] = \sum_{j=1}^{M} \frac{P[X_{n+1} = j, X_n = i]}{P[X_n = i]}$$

$$= \frac{\sum_{j=1}^{M} P[X_{n+1} = j, X_n = i]}{P[X_n = i]} = \frac{P[X_n = i]}{P[X_n = i]} = 1$$

1.3 Question 3 : Distribution of Step n

If in a Markov Sequence $\{X_n\}_{n=0}^{\infty}$, distribution of X_n is denoted by λ_n , Prove that :

$$\lambda_n = \lambda P^n$$

Proof 1 : By Using Matematical Induction .

N = 0 : $\lambda_0 = \lambda$ (Base Case)

Suppose that foe N = k is true , and prove for N = k + 1

N = k : $\lambda_k = \lambda P^k$

and now let's prove for N = k + 1 :

N = k + 1 : $\lambda_{k+1} = \lambda P^{k+1} = (\lambda P^k) \cdot P = \lambda_k P$

It's obviouse that $\lambda_{k+1} = \lambda_k \cdot P$ is true , because it's the Markov chain's property that each step is transformed to the next one, when it's multiplied by P .

Proof 2 : We know that :

$$\lambda_n = \lambda_{n-1} \cdot P$$

So we have :

$$\lambda_n = \lambda_{n-1} \cdot P = (\lambda_{n-2} \cdot P) \cdot P = \cdots = \lambda_0 \cdot P^n$$

1.4 Question 4 : Higher Order Markov Chain

1.4.1 Markov Chain Of 2_{nd} Order

If for a sequence $[X_n]_{n=0}^{\infty}$ we have :

$$P[X_{n+1}|X_0, X_1, \cdots, X_{n-1}, X_n] = P[X_{n+1}|X_n, X_{n-1}] \qquad \text{(Second Order Markov Chain)}$$

Prove that it can be written as a Morkov Chain (First Order Markov Chain) .

Proof : Let's define $Y_n := (X_n, X_{n+1})$

So Equation :

$$P[X_{n+1}|X_0, X_1, \cdots, X_{n-1}, X_n] = P[X_{n+1}|X_n, X_{n-1}]$$

Can be Witten as :

$$P[X_{n+1}|X_0, X_1, \cdots, X_{n-1}, X_n] = P[Y_n|Y_{n-1}]$$

We can show that :

$$P[Y_n|Y_{n-1}] = P[(X_{n+1}, X_n)|(X_n, X_{n-1})] = P[X_{n+1}|X_n, X_{n-1}]$$

1.4.2 Markov Chain Of K_{th} Order :

If for a sequence $[X_n]_{n=0}^{\infty}$ we have :

$$P[X_{n+1}|X_0, X_1, \cdots, X_{n-1}, X_n] = P[X_{n+1}|X_n, X_{n-1}, \cdots, X_{n-k}]$$

Prove that it can be written as a Morkov Chain (First Order Markov Chain) .

Proof : Let's define $Y_n := (X_{n-k+1}, X_{n-k+2}, \cdots, X_{n+1})$

So the Equation :

$$P[X_{n+1}|X_0, X_1, \cdots, X_{n-1}, X_n] = P[X_{n+1}|X_n, X_{n-1}, \cdots, X_{n-k}]$$

Can be written as :

$$P[X_{n+1}|X_0, X_1, \cdots, X_{n-1}, X_n] = P[Y_n|Y_{n-1}]$$

And we show that :

$$P[Y_n|Y_{n-1}] = P[(X_{n+1}, X_n, \cdots, X_{n-k+1})|(X_n, X_{n-1}, \cdots, X_{n-k})]$$

$$= P[X_{n+1}|X_n, X_{n-1}, \cdots, X_{n-k}]$$

1.5 Question 5

Prove that sequence $[X_n]_{n=0}^{\infty}$, is a Markov Chain with initial distribution of λ and transition matrix of $P = [p_{ij}]_{M\times M}$, If and only if ($\forall\, n$ and $\{i_N\}_{N=0}^{n} \in\ \chi$) we have :

$$P[X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n] = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_{n-1} i_n}$$

Proof : Suppose we observe a finite realization of the discrete Markov chain and want to compute the probability of this random event:

Using Conditional Probability :

$$P[X = x\,, Y = y] = P[X = x|Y = y] \cdot P[Y = y]$$

We have :

$$P[X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1, X_0 = i_0]$$

$$= P[X_n = i_n | X_{n-1} = i_{n-1}, \cdots, X_1 = i_1, X_0 = i_0] \cdot P[X_{n-1} = i_{n-1}, \cdots, X_1 = i_1, X_0 = i_0]$$

$$= p_{i_{n-1}i_n} \cdot P[X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}, \cdots, X_1 = i_1, X_0 = i_0] \cdot P[X_{n-2} = i_{n-2}, \cdots, X_1 = i_1, X_0 = i_0]$$

$$= \vdots$$

$$= p_{i_{n-1}i_n} p_{i_{n-2}i_{n-1}} \cdots p_{i_0 i_1} P[X_0 = i_0]$$

$$= \lambda_{i_0} \cdot p_{i_0 i_1} \cdot p_{i_1 i_2} \cdots p_{i_{n-2}i_{n-1}} \cdot p_{i_{n-1}i_n}$$

1.6 Question 6

Draw the state diagram for the below transition matrix :

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Here is the State Diagram :

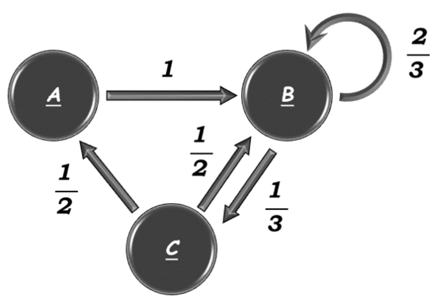

Figure 3: State diagram of matrix P

1.7 Question 7 : The Chapman–Kolmogorov equations

First , Let's mention a proposition that , In general, we define the n-stage transition probabilities, denoted as $P_{ij}^{(n)}$, by

$$p_{ij}(m, m+n) = P[X_{m+n} = j, X_m = i] = p_{ij}^{(n)}$$

For a Markov chain, P_{ij} represents the probability that a system in state i will enter state j at the **next transition**. We can also define the **n-stage transition** probability $P_{ij}^{(n)}$ that a system presently in state i will be in state j after n additional transitions.

Let's Start by the proof for the Equation

$$P[X_{m+n} = j, X_m = i] = P[X_n = j, X_0 = i] = p_{ij}^{(n)}$$

Proposition 7.1 :

Proof :

$$P[X_{m+n} = j | X_m = i] =$$

$$= \sum_{i_{m+1}, i_{m+2}, \cdots, i_{m+n-1}} P[X_{m+n} = j | X_{m+n-1} = i_{m+n-1}] \times \cdots \times P[X_{m+1} = i_{m+1} | X_m = i_m]$$

$$= \sum_{i_{m+1}, i_{m+2}, \cdots, i_{m+n-1}} P[X_n = j | X_{n-1} = i_{m+n-1}] \times \cdots \times P[X_1 = i_{m+1} | X_0 = i_m]$$

$$= P[X_n = j | X_0 = i] = p_{ij}^{(n)}$$

Now , Let's use the **Proposition 7.1** to prove that :

$$p_{ij}^{(n+r)} = p_{ij}(m, m+n+r) = \sum_k p_{ik}(m, m+n) \cdot p_{kj}(m+n, m+n+r)$$

Proof :

$$p_{ij}^{(n+r)} = P[X_{n+r} = j | X_0 = i] = \sum_{k \in S} P[X_{n+r} = j, X_n = k | X_0 = i]$$

$$= \sum_{k \in S} P[X_{n+r} = j | X_n = k, X_0 = i] = \sum_{k \in S} P[X_{n+r} = j | X_n = k] \cdot P[X_n = k | X_0 = i]$$

$$= \sum_{k \in S} p_{ik}^{(n)} \cdot p_{kj}^{(r)} = \sum_{k \in S} p_{ik}(m, m+n) \cdot p_{kj}(m+n, m+n+r)$$

Now if we know that $\rightarrow P(m, m+n) = [p_{ij}(m, m+n)]_{M \times M}$

Prove that $\Rightarrow P(m, m+n) = P^n \qquad (*)$

Proof :

First, We know that the i_{th} row and j_{th} column of matrix P^n is denoted by $p_{ij}^{(n)}$.

Second, We know that each member of matrix $P(m, m+n)$ is denoted by $p_{ij}(m, m+n)$.

For the proof , let's show that both matrices are the same in the i_{th} row and j_{th} column .

So we have :

$$p_{ij}(m, m+n) = P[X_{m+n} = j | X_m = i] =$$

$$= \sum_{i_{m+1}, i_{m+2}, \cdots, i_{m+n-1}} P[X_{m+n} = j | X_{m+n-1} = i_{m+n-1}] \times \cdots \times P[X_{m+1} = i_{m+1} | X_m = i_m]$$

$$= \sum_{i_{m+1}, i_{m+2}, \cdots, i_{m+n-1}} P[X_n = j | X_{n-1} = i_{m+n-1}] \times \cdots \times P[X_1 = i_{m+1} | X_0 = i_m]$$

$$= P[X_n = j | X_0 = i] = p_{ij}^{(n)}$$

Then since : $p_{ij}(m, m+n) = p_{ij}^{(n)} \Longrightarrow P(m, m+n) = P^n$

1.8 Question 8 : Classifiction of States

- Accessible: State j is accessible from state i if $P_{ij}^{(n)} > 0 \qquad (\forall n \geq 0)$

$$P_{ij}^{(n)} > 0 \qquad (\forall n \geq 0) \qquad \Longleftrightarrow \qquad i \to j$$

Now we see that $(i \to j)$ is equivalent to : $P[\text{ ever enter } j \,|\, \text{start in } i] > 0$

Proposition 8.1 :

$$P[\text{ ever enter } j \,|\, \text{start in } i]$$

$$= P[\cup_{n=0}^{\infty}\{X_n = j\}|X_0 = i]$$

$$\leq \sum_{n=0}^{\infty} P[X_n = j|X_0 = i]$$

$$= \sum_{n=0}^{\infty} p_{ij}^{(n)}$$

Hence if $p_{ij}^{(n)} = 0$ (for all n) , $P[\text{ ever enter } j \,|\, \text{start in } i] = 0$.

On the other hand,

Proposition 8.2 :

$$P[\text{ ever enter } j \,|\, \text{start in } i]$$

$$= P[\cup_{n=0}^{\infty}\{X_n = j\}|X_0 = i]$$

$$\geq P[X_n = j|X_0 = i] \qquad \text{(for any n)}$$

$$= p_{ij}^{(n)}$$

If $p_{ij}^{(n)} > 0$ for some n, $P[\text{ ever enter } j \,|\, \text{start in } i] \geq p_{ij}^{(n)} > 0$.

1.9 Question 9

First , Let's define **Equivalance relation**

- An equivalence relation "$\sim$" is a binary relation between elements of a set satisfying

  1. Reflexivity: $i \sim j$ for all i
  2. Symmetry: $i \sim j \Longrightarrow j \sim i$
  3. Transitivity: $i \sim j, j \sim k \Longrightarrow i \sim k$

Proposition 9.1 : Communication of states is an equivalence relation.

Proof : Reflexivity and symmetry are clear. To prove transitivity, let $i \leftrightarrow j$ and $j \leftrightarrow k$. We then want to show that $i \leftrightarrow k$.

- Note that $i \to j$ if and only if the transition diagram contains a path from i to j.

So there is a path from i to j and from j to k. Concatenate the two to obtain a path from i to k, which testifies to the fact that $i \to k$.

Analogously, we have $k \to i$ and conclude that $i \leftrightarrow k$.

1.10 Question 10

Find the communication classes of a markov chain with this transition matrix :

$$P = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

Which classes are close ?

Let's draw the state diagram of the matrix :

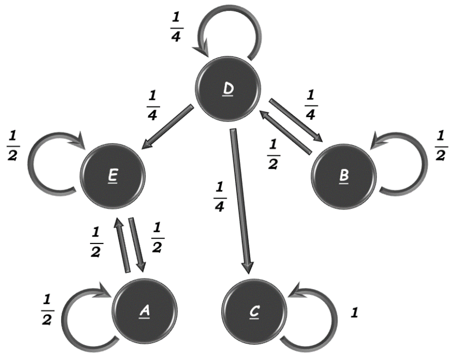

Figure 4: State diagram

The communication classes are : {A} , {B}, {C}, {D}, {E} , {A,E} , {B,D}

The close communication classes are : {C}, {A,E}

1.11 Question 11 : Random walk (one dimentional)

There's a man walking on a line of integer numbers, steps forward with probability of p and steps back with $1 - p$, and now we're gonna find the period of the states .(The start point is at $X_0 = 0$).

The mathematically modeled form of the problem is :

$$P[x_{n+1} = i + 1 | X_n = i] = p \qquad \text{and} \qquad P[x_{n-1} = i - 1 | X_n = i] = p$$

The states of this markov chain are integer numbers. So, we will have **odd** or **even** periods.

• Define $p_{ii}^{(n)}$ is the probability of starting at state i and turn back to state i after n steps .

It's **obvious** that if you are in the state $\boldsymbol{i}$, it's impossible to reach state $\boldsymbol{i}$ again after an odd number of steps .

E.g, suppose that the man is on the state *0* , it's obvious that there's noway he can turn back to the *0* state by moving odd (1,3,5,$\cdots$) steps; So, we have :

$$p_{ii}^{(n)} = 0 \qquad , (\forall\ \boldsymbol{n} = 2k + 1\ ;\ \boldsymbol{k} = 0, 1, 2, \cdots)$$

We want to show that for all $\boldsymbol{i}$ states on the line ,

$$p_{ii}^{(n)} > 0 \qquad , \ (\forall \ \boldsymbol{n} = 2k \ ; \ \boldsymbol{k} = 0, 1, 2, \cdots)$$

As we know that moves with even steps can be modeled by 2-step moves, Let's Prove that :

$$p_{ii}^{(2)} > 0 \qquad , \ (\boldsymbol{n} = 2)$$

Proof :

$$p_{ii}^{(2)} = P[X_2 = i | X_0 = i] = \sum_k P[X_2 = i, X_1 = k | X_0 = i]$$

$$= \sum_k P[X_2 = i | X_1 = k, X_0 = i] \cdot P[X_1 = k | X_0 = i]$$

$$= \sum_k P[X_2 = i | X_1 = k] \cdot P[X_1 = k | X_0 = i]$$

$$= P[X_2 = i | X_1 = i-1] \cdot P[X_1 = i-1 | X_0 = i] + P[X_2 = i | X_1 = i+1] \cdot P[X_1 = i+1 | X_0 = i]$$

$$= p(1-p) + (1-p)p = 2p(1-p) > 0$$

So we have : $\qquad (d_i = \gcd\{\boldsymbol{n} = 2k \geq 1; \boldsymbol{k} = 1, 2, \cdots\} = 2) \qquad$ (even steps)

1.12 Question 12

Period is a communication class property. Namely, $i \leftrightarrow j \Longrightarrow d_i = d_j$

Let $p_{ij}^{(k)}$ denote the k-step transition probability from state i to j and let d_i denote i's period.

Suppose that i and j communicating and $\alpha = p_{ij}^{(k)} > 0$ and $\beta = p_{ji}^{(l)}$ for some k and l. Then by Chapman-Kolmogrov equations :

$$p_{ii}^{(k+l)} = \sum_k p_{ik}^{(k)} p_{ki}^{(l)} \quad \geq \quad p_{ij}^{(k)} p_{ji}^{(l)} = \alpha\beta > 0$$

So , $(k+l)$ is multiple of $d_i \Longrightarrow d_i|(k+l)$

Suppose now that $m \in \{n \ ; \ p_{jj}^{(n)} > 0\}$ then again using Chapman-Kolmogrov equations we have that $p_{ii}^{(k+l+m)} > 0$ (you go from i to j, go off somewhere and return to j, and then return to i), so $d_i|(k+l+m)$ and so we get that $d_i|m$. Since $d_i|m$ and $d_i|(k+l+m)$, so $d_j \geq d_i$. Now exchanging the roles of i and j we get that $d_i \geq d_j$.

So we conclude that $d_i = d_j$.

1.13 Question 13 : Hitting probabilities

Let's prove that the vector of hitting probabilities, $h^A = [h_1^A, h_2^A, \cdots, h_M^A]$, minimal answer of the following system of linear equations:

$$h_i^A = \begin{cases} 1 & \forall i \in A \\ \sum_{j \in X} p_{ij} h_j^A & \forall i \notin A \end{cases} \tag{1}$$

Proof : We consider to cases seperately . If $X = i \in A$, then $H^A = 0$ and $h_i^A = 1$.

On the other hand, if $X = i \notin A$, then $H^A \geq 1$, so by the Markov property :

$$P_i[H^A < \infty | X_1 = j] = P[H^A < \infty] = h_j^A$$

Now, Total probability theorem implies that :

$$h_i^A = P_i[H^A < \infty] = \sum_{j \in S} P_i[H^A < \infty | X_1 = j] P_i[X_1 = j] = \sum_{j \in S} p_{ij} h_j^A$$

1.14 Question 14

Let's prove that the vector of mean hitting times, k^A ,is the minimal non-negative solution to the following system of linear equations :

$$k_i^A = \begin{cases} 0 & \forall i \in A \\ 1 + \sum_{j \in \mathcal{X}} p_{ij} k_j^A & \forall i \notin A \end{cases} \tag{2}$$

Proof : We consider to cases seperately . If $X = i \in A$, then $H^A = 0$ and $k_i^A = 0$.

On the other hand, if $X = i \notin A$, then $H^A \geq 1$, so by the Markov property :

$$E_i[H^A | X_1 = j] = 1 + E_j[H^A]$$

consequently, by the partition theorem for the expectations,

$$k_i^A = E_i[H^A] = \sum_{j \in S} E_i[H^A | X_1 = j] P_i[X_1 = j] = \sum_{j \in S} p_{ij} k_j^A$$

1.15 Question 15 : Example

According to the transition matrix :

1. Solution : Put $h_i = P_i[H^{\{4\}} < \infty]$. Clearly $h_1 = 0$ and $h_4 = 1$ and by using markov property we have :

$$h_2 = \tfrac{1}{2}(h_1 + h_3) \qquad h_3 = \tfrac{1}{2}(h_2 + h_4)$$

As the result : $h_2 = \frac{1}{2} h_3 = \frac{1}{2}(\frac{1}{2} h_2 + \frac{1}{2})$, that is $h_2 = \frac{1}{3}$ and $h_3 = \frac{2}{3}$.

2. Solution : Put $k_i = E_i[h^{\{1,4\}}]$. Clearly, $k_1 = k_2 = 0$ and by using Markov chain property we have : $k_2 = 1 + \frac{1}{2}(k_1 + k_3)$ $\qquad$ $k_3 = 1 + \frac{1}{2}(k_2 + k_4)$

As the result : $k_2 = 1 + \frac{1}{2} k_3 = 1 + \frac{1}{2}(\frac{1}{2} k_2)$, that is $k_2 = k_3 = 2$.

1.17 Question 17 :

Let's More generally prove that :

$$\text{State } \mathbf{i} \text{ is } \mathbf{recurrent} \iff \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$$

$$\text{State } \mathbf{i} \text{ is } \mathbf{transient} \iff \sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$$

Proof : First, Let's remind some defenitions :

Definition 17.1 : A state $i \in S$ is recurrent if :

$$P[\exists\ n \geq 1 : X_n = i | X_0 = i] = 1$$

Definition 17.2 : A state $i \in S$ is recurrent if :

$$P[\exists\ n \geq 1 : X_n = i | X_0 = i] = 0$$

Definition 17.2 : A state $i \in S$, Denote f_{ii}, the probability that a Markov chain which starts at i returns to i at least once, that is :

$$f_i = P[\exists\ n \in \mathbb{N} : X_n = i]$$

Then ,

1. State i is recurrent if and only if $f_i = 1$

2. State i is transient if and only if $f_i < 1$

Now let's have the main question :

Let the Markov chain start at state i. Consider the random variable

$$V_i := \sum_{n=1}^{\infty} I\{X_n = i\}$$

Which counts the number of returns of the Markov chain to state i. Note that the random variable V_i can take the value $+\infty$. Then,

$$P[B_k] = P[V_i \geq k] = f_i^k \qquad (\forall\ k \in \mathbb{N})$$

Thus, the expectation of V_i can be computed as follows:

$$\mathbb{E}_i[V_i] = \sum_{k=1}^{\infty} P_i[V_i \geq k] = \sum_{k=1}^{\infty} f_i^k \tag{17.1}$$

On the other hand ,

$$\mathbb{E}_i[V_i] = \mathbb{E}_i\left[\sum_{n=1}^{\infty} I\{X_n = i\}\right] = \sum_{n=1}^{\infty} \mathbb{E}_i[I\{X_n = i\}] = \sum_{n=1}^{\infty} p_{ii}^{(n)} \tag{17.2}$$

Case 1 : Assume that state i is recurrent. Then, By definition 17.3 $\Rightarrow f_i = 1$.

It follows by 17.1 that, $\mathbb{E}_i[V_i] = \infty$.

(In fact, $P_i[V_i = +\infty] = 1$ since $P[v_i \geq k] = 1$ for every $k \in \mathbb{N}$) .

Hence, by 17.2, $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$

Case 2 : Assume that state i is transient. Then, By definition 17.3 ,$f_i < 1$.Thus, By 17.1, $\mathbb{E}_i[V_i] < \infty$. Hence, By 17.2, $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$.

1.20 Question 20 : Recurrence and transience of random walks

A simple random walk on $\mathbb{Z}$ is a Markov chain with state space $S = \mathbb{Z}$ andtransition probabilities :

$$p_{i,i+1} = p \qquad , \qquad p_{i,i-1} = 1 - p \qquad i \in \mathbb{Z}$$

So, from every state the random walk goes one step to the forward with probability p, or one step back with probability $1-p$;

Look at Figure 5 that simulates **Random walk** for 200 steps on $\mathbb{Z}$ with $p = \frac{1}{2}$.

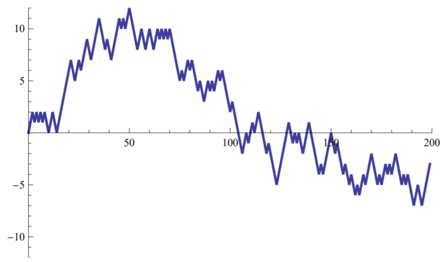

Figure 5: Random Walk Simulation with $p = \frac{1}{2}$

Theorem 20.1 : If $p = \frac{1}{2}$, then any state of the simple random walk is recurrent.

If $p \neq \frac{1}{2}$ then any state is transient.

References :

1 . Probability - Random Variables and Stochastic Processes.
(by:Athanasios Papoulis)

2 . A First Course in Probability (9th Edition). (by: Sheldon Ross)

3 . Cambridge University Statistics Courses :
(https://www.training.cam.ac.uk/course/ucs-stats_diy)

4 . Discrete Probability Models and Methods: Probability on Graphs and Trees, Markov Chains and Random Fields, Entropy and Coding .
(by : Pierre Brémaud) .