First Step Analysis. Extended Example

These notes provide two solutions to a problem stated below and discussed in lectures (Sections 1, 2). The difference between these solutions will be discussed in Section 3 where also several examples of problems which can be solved by these methods are given. **Statement of the Problem.** Which will take fewer flips, on average: successively flipping a fair coin until the pattern HHT appears, that is, you observe two successive heads followed by a tails; or successively flipping the same coin until the pattern HTH appears?

1 The "universal" approach.

First of all let us agree about the following notations: we write 1 for H and 0 for T. (The letter T will denote (as before) the time of absorbtion of the MC.) In these notation, tosses are presented by a sequence of independent random variables ξ_0, $\xi_1, ..., \xi_n, ...$ with $P\{\xi_n = 1\} = \frac{1}{2}$ and $P\{\xi_n = 0\} = \frac{1}{2}$. Then by definition X_n is a random variable which assigns to n the pattern $(\xi_n, \xi_{n+1}, \xi_{n+2})$. We thus could simply write $X_0 = (\xi_0, \xi_1, \xi_2)$, $X_1 = (\xi_1, \xi_2, \xi_3)$, and so on. However, it is more convenient to list all possible patterns of such triples in some order and then to work with their order numbers (the analysis does not depend on the choice of the order). There are 8 such patterns and we shall order them as follows:

$$1 \leftrightarrow (101),\ 2 \leftrightarrow (100),\ 3 \leftrightarrow (001),\ 4 \leftrightarrow (000),\ 5 \leftrightarrow (011),\ 6 \leftrightarrow (010),\ 7 \leftrightarrow (111),\ 8 \leftrightarrow (110).$$

(If you have a difficulty at this point then it may be instructive to understand the following example. If our random sequence shows up as 01001000101 then $X_0 = 6,\ X_1 = 2,\ X_2 = 3,\ X_3 = 6,\ X_4 = 2,\ X_5 = 4,\ X_6 = 3,\ X_7 = 6,\ X_8 = 1$.)

We remark that the sequence X_n is a MC because the distribution of X_{n+1} is uniquely defined if X_n is given and is independent of the past values of this sequence.

Because the eight patterns are equally likely to be observed on the first three tosses we set $P(X_0 = j) = 1/8$, $j = 1, 2, \dots 8$ which is thus the initial probability distribution of our MC.

We now have to consider two Markov chains describing (a) the process of flipping the coin until $X_T = 8$ appears (in other words the pattern (110) appears); and (b) the process of flipping the coin until $X_T = 1$ appears (the pattern (101) appears.

(a) State 8 is absorbing. Accordingly the (one-step) transition probability matrix is

$$\mathbb{P} = \left\| \begin{array}{cccccccc} 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right\|$$

and the mean number of tosses required to finish the game is $E(T) + 3$.

Denote $E(T|X_0 = j) = v_j$, $j = 1, \ldots, 8$. Obviously $v_8 = 0$. By the first step analysis,

$$\begin{aligned} v_1 &= 1 + \frac{1}{2}v_5 + \frac{1}{2}v_6 \\ v_2 &= 1 + \frac{1}{2}v_3 + \frac{1}{2}v_4 \\ v_3 &= 1 + \frac{1}{2}v_5 + \frac{1}{2}v_6 \\ v_4 &= 1 + \frac{1}{2}v_3 + \frac{1}{2}v_4 \\ v_5 &= 1 + \frac{1}{2}v_7 \\ v_6 &= 1 + \frac{1}{2}v_1 + \frac{1}{2}v_2 \\ v_7 &= 1 + \frac{1}{2}v_7. \end{aligned}$$

The only solution to these equations is $v_1 = v_3 = 6$, $v_2 = v_4 = v_6 = 8$, $v_5 = v_7 = 2$. By the law of total probability,

$$\begin{aligned} E(T) &= \sum_{j=1}^{8} E(T|X_0 = j) \times P(X_0 = j) \\ &= \frac{1}{8}\sum_{j=0}^{8} E(T|X_0 = j) = \frac{1}{8}(v_1 + v_2 + v_3 + v_4 + v_5 + v_6 + v_7 + v_8) = 5 \end{aligned}$$

and the mean number of tosses required to finish the game is 5+3=8.

(b) Now state 1 corresponding to (101) is absorbing. Accordingly the one-step transition probability matrix is

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

and again the mean number of tosses required to finish the game is $E(T) + 3$.

Denote $E(T|X_0 = j) = v_j$, $j = 1, \ldots, 8$. Obviously now $v_1 = 0$ and $v_8 \neq 0$. By the first step analysis (see the derivation of equations in your lecture notes or in those posted on the web) we have:

$$\begin{aligned} v_2 &= 1 + \frac{1}{2}v_3 + \frac{1}{2}v_4 \\ v_3 &= 1 + \frac{1}{2}v_5 + \frac{1}{2}v_6 \\ v_4 &= 1 + \frac{1}{2}v_3 + \frac{1}{2}v_4 \\ v_5 &= 1 + \frac{1}{2}v_7 + \frac{1}{2}v_8 \\ v_6 &= 1 + \frac{1}{2}v_2 \\ v_7 &= 1 + \frac{1}{2}v_7 + \frac{1}{2}v_8 \\ v_8 &= 1 + \frac{1}{2}v_2. \end{aligned}$$

the only solution to these equations is $v_2 = v_4 = 10$, $v_3 = v_5 = v_7 = 8$, $v_6 = v_8 = 6$. By the law of total probability

$$\begin{aligned} E(T) &= \sum_{j=0}^{7} E(T|X_0 = j) \times P(X_0 = j) \\ &= \frac{1}{8}\sum_{j=0}^{8} E(T|X_0 = j) = \frac{1}{8}(v_1 + v_2 + v_3 + v_4 + v_5 + v_6 + v_7 + v_8) = 7 \end{aligned}$$

and the mean number of tosses required to finish the game is 7+3=10.

Flipping a fair coin until the pattern HTH appears will take, on average, 10 flips, 2 more than for HHT.

2 The "problem tailored" approach

We shall only solve the case (b) (the other one can be done similarly). As before we write 1 for H and 0 for T and the sequence of tosses is presented by a sequence of independent random variables $\xi_0, \xi_1, ..., \xi_n, ...$ with $P\{\xi_n = 1\} = \frac{1}{2}$ and $P\{\xi_n = 0\} = \frac{1}{2}$.

Consider the set $S = \{\mathbb{O}, (1), (10), (101)\}$. We shall set up a MC with the state space S. In order to do that, imagine that you are a player and you are waiting for (101) to appear. While waiting, you are recording the results in the following way.

You start by recording $\mathbb{O}$ which marks the beginning of the game. After that you are supposed to make a record after each tossing of the coin.

If 0 appears at the first tossing then you record $\mathbb{O}$ which means that you have not got any closer to the result and are at the beginning of your "waiting duty".

When you observe 1 for the first time you record it as (1) since this is the beginning of (101).

If you have just recorded (1) and you observe a 0 at the next tossing then your next record is (10) (and you are just one step from the (101)!); but if (1) is followed by 1 then your new record is once again (1).

If you have just recorded (10) and you obtain 0 at the next tossing then your next record is $\mathbb{O}$ meaning again that you returned to the beginning of the game.

Note now that your sequence of records which we shall call Y_n is a Markov chain with values in S. As usual we prefer to work simpler notations for the states of the chain and therefore we establish the following one-to-one correspondence

$$1 \leftrightarrow \mathbb{O},\ 2 \leftrightarrow (1),\ 3 \leftrightarrow (10),\ 4 \leftrightarrow (101).$$

Example. Suppose that 01100101 is a sequence of observed values (tosses). Then your record is $\mathbb{O}$, $\mathbb{O}$, (1), (1), (10), $\mathbb{O}$, (1), (10), (101), or in the more covenient notation, $Y_0 = 1$, $Y_1 = 1$, $Y_2 = 2$, $Y_3 = 2$, $Y_4 = 3$, $Y_5 = \mathbb{O}$, $Y_6 = 2$, $Y_7 = 3$, $Y_8 = 4$.

The transition matrix of the chain is now easily found. Indeed it is clear that

$$p_{11} = p_{12} = \frac{1}{2},\ p_{22} = p_{23} = \frac{1}{2},\ p_{31} = p_{34} = \frac{1}{2},\ p_{44} = 1$$

and all other one-step transition probabilities are equal to 0. Hence

$$\mathbb{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

As usual we set $v_i = E(T|X_0 = i)$. The equations are of the same kind as before:

$$\begin{aligned} v_1 &= 1 + \frac{1}{2}v_1 + \frac{1}{2}v_2 \\ v_2 &= 1 + \frac{1}{2}v_2 + \frac{1}{2}v_3 \\ v_3 &= 1 + \frac{1}{2}v_1 + \frac{1}{2}v_4 \\ v_4 &= 0 \end{aligned}$$

Solving these these equations gives $v_1 = 10$, $v_2 = 8$, $v_3 = 6$. The answer is 10 (as we already know).

3 What is the difference between these solutions?

The **first approach** can be used to solve a wide variety of problems. For instance, essentially the same Markov chain can be used in order to answer the following questions:

1. A fair coin is being flipped repeatedly. What is the probability that the sequence HTH will be seen before THH?
In terms of the MC this question sounds: given that (101) and (011) are absorbing states of the MC defined in Section 1, what is the probability that the chain will be absorbed by (101)?

2. You are asked whether you would want to take part in the following game. A fair coin is flipped repeatedly until either HHH or HTT shows up for the first time. Every time when a sequence TTT shows up, you are given £1. Every time HTH shows up, you pay £1.

To answer this and similar questions you can again use the same MC. If the mean value of your gain is positive then the game is worth playing. The way to find this mean value was discussed in the lectures. A complete account of this discussion is contained in example 4 explained in the notes concerning the FSA (which are on the web).

The **second approach** allows one to set up a Markov chain with a state space S designed to solve a particular problem. That is why the size $\#(S)$ of the state space can be much smaller than in the general case, e.g. in our case these sizes are 8 and 4. This difference may be much more serious; for instance in problem 3 from the CW2 the general approach would require a MC with a 36×36 transition matrix whereas the "tailored' approach requires a 5×5 transition matrix.

On the other hand the necessity of setting separate Markov chains for each problem in a series of similar problems may be a disadvantage.

Exercise. Set up a MC which would "suit" the following problem. A fair die is rolled until the sum of three consecutive rolls is 4 for the first time. What is the mean number of rolls in this game?