

M362M: Introduction to Stochastic Processes

Practice Midterm III

Name: Lee Wei LuanUT EID: w122963

Instructions:

- This exam contains 4 questions on 9 pages.
- Show all your work to receive full credit. Final answers without justification will get no credit.
- You may use your notes. However, electronic devices of any type are not allowed.
- Bathroom breaks are not allowed, except on emergency situations.

1. Consider the following R script:

```

1 # state space
2 S = c(1,2,3,4,5)
3 # transition matrix
4 P <- matrix(c(
5   1/2, 0, 0, 0, 1/2,
6   0, 1/2, 0, 1/2, 0,
7   1/4, 1/4, 1/2, 0, 0,
8   0, 1/2, 0, 1/2, 0,
9   1/2, 0, 0, 0, 1/2
10 ), nrow = 5, byrow = TRUE)
11
12 # simulate the next position of the chain
13 draw_next = function(s) {
14   sample(S, prob = P[s, ], size = 1)
15 }
16 # simulate a single trajectory of length T from the initial distribution
17 single_trajectory = function(initial_distribution, T) {
18   path = numeric(T)
19   path[1] = sample(S, prob = initial_distribution, size = 1)
20   for (n in 1:(T-1)) {
21     path[n+1] = draw_next(path[n])
22   }
23   return(path)
24 }
25 # simulate the entire chain
26 simulate_chain = function(initial_distribution, nsim, T) {
27   initial_state =
28     data.frame(t(replicate(
29       nsim, single_trajectory(initial_distribution, T))
30     ))
31 }

```

```

1/
2 init_dist=c(0,1,0,0,0)
3 nsim=1000
4 T=100
5 df = simulate_chain(init_dist,nsim,T)

```

Listing 1: Monte Carlo Sampling in R

Give a possible output of

(a) `single_trajectory(c(0,1,0,0,0),10)`

2 2 2 2 2 2 2 2 2 2.

Consider the function

```

1 tau_1=function(x){
2   for (k in 1:length(x)){
3     if (x[k]==1){
4       return(k)
5     }
6   }
7   return(length(x))
8 }
9 tau_4=function(x){
10  for (k in 1:length(x)){
11    if (x[k]==4){
12      return(k)
13    }
14  }
15  return(length(x))
16 }
17 }

```

```
18 T_1=apply(df,1,tau_1)
19 T_4=apply(df,1,tau_4)
```

Write down the expected output of

(b) `mean(T_1)`

Note that $\{2, 4\}$ is a closed set of states for the Markov Chain.

Therefore we will never have $X_k = 1$ for $k = 1, \dots, 100$ if we start with $X_1 = 2$.

Thus we expect to see 100.

(c) `mean(T_4)`

Let ν be the stopping time

$$\nu = \min\left(\inf\{n \in \mathbb{N} : X_n = 4\}, 100\right).$$

$$\mathbb{E}[\nu] = \sum_{k=1}^{100} k \cdot P(\nu = k).$$

$$\begin{aligned}
 &= 1 \cdot 0 + \sum_{k=2}^{99} k \cdot \left(\frac{1}{2}\right)^{k-1} + 100 \cdot \left[\left(\frac{1}{2}\right)^{99} + \left(\frac{1}{2}\right)^{99} \right] \\
 &= 0 + \frac{1}{1 - \frac{1}{2}} + \frac{\left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^{99}}{\left(1 - \frac{1}{2}\right)^2} - \frac{99 \cdot \left(\frac{1}{2}\right)^{99}}{1 - \frac{1}{2}} + 100 \cdot \left(\frac{1}{2}\right)^{98} \\
 &= 3 - \left(\frac{1}{2}\right)^{97} + \left(\frac{1}{2}\right)^{98}
 \end{aligned}$$

$X_1 = \dots = X_{100} = 2$ $X_1 = \dots = X_{99} = 2, X_{100} = 4$

2. Consider the following R script:

```

1 # state space
2 S = c(1,2,3)
3 # transition matrix
4 P <- matrix(c(
5   1/2, 1/2, 0,
6   1/5, 2/5, 2/5,
7   1/5, 1/5, 3/5
8 ), nrow = 3, byrow = TRUE)
9
10 # simulate the next position of the chain
11 draw_next = function(s) {
12   sample(S, prob = P[s, ], size = 1)
13 }
14 # simulate a single trajectory of length T from the initial distribution
15 single_trajectory = function(initial_distribution, T) {
16   path = numeric(T)
17   path[1] = sample(S, prob = initial_distribution, size = 1)
18   for (n in 1:(T-1)) {
19     path[n+1] = draw_next(path[n])
20   }
21   return(path)
22 }
23 # simulate the entire chain
24 simulate_chain = function(initial_distribution, nsim, T) {
25   initial_state =
26     data.frame(t(replicate(
27       nsim, single_trajectory(initial_distribution, T))
28     ))
29 }
30 init_dist = c(1, 0, 0)
31 nsim = 1000
32 T = 1000
33 df = simulate_chain(init_dist, nsim, T)

```

Listing 2: Monte Carlo Sampling in R

What do you expect the output of

(a) `mean(df$X2==1)`

$$P(X_2 = 1 \mid X_1 = 1) = \frac{1}{2}.$$

(b) `mean(df$X1000==1)`

Since this Markov chain is irreducible, we have a stationary distribution $\pi^* = (\pi_1^*, \pi_2^*, \pi_3^*)$ where $\pi_i^* = \lim_{n \rightarrow +\infty} P(X_n = i)$, we expect to see π_1^* .

We know that

$$\pi^* P = \pi^* \text{ and } \sum_{i=1}^3 \pi_i^* = 1, \pi_i^* \geq 0$$

$$\text{So } \begin{cases} \frac{1}{2}\pi_1^* + \frac{1}{5}\pi_2^* + \frac{1}{5}\pi_3^* = \pi_1^* \\ \frac{1}{2}\pi_1^* + \frac{2}{5}\pi_2^* + \frac{1}{5}\pi_3^* = \pi_2^* \\ \frac{2}{5}\pi_2^* + \frac{3}{5}\pi_3^* = \pi_3^* \end{cases} \Rightarrow \begin{cases} \pi_2^* = \frac{5}{4}\pi_1^* \\ \pi_3^* = \frac{5}{4}\pi_1^* \end{cases}$$

$$\therefore \pi_1^* = \frac{2}{7}, \pi_2^* = \frac{5}{14}, \pi_3^* = \frac{5}{14}$$

(c) `mean(df$X1000)`

Suppose X_{1000} follow the distribution π^*

Then

$$\begin{aligned} E[X_{1000}] &= 1 \cdot \frac{2}{7} + 2 \cdot \frac{5}{14} + 3 \cdot \frac{5}{14} \\ &= \frac{29}{14} \\ &= 2.0714. \end{aligned}$$

3. Consider the following R script:

```

1 # state space
2 S = c(1,2,3,4)
3 # transition matrix
4 P <- matrix(c(
5   1/3, 1/3, 1/3, 0, # Transitions from state 1
6   0,   1,   0,   0, # Transitions from state 2
7   0,   0,   1/2, 1/2, # Transitions from state 3
8   0,   0,   0,   1, # Transitions from state 4
9 ), nrow = 4, byrow = TRUE)
10
11 # simulate the next position of the chain
12 draw_next = function(s) {
13   sample(S, prob = P[s, ], size = 1)
14 }
15 # simulate a single trajectory of length T from the initial distribution
16 single_trajectory = function(initial_distribution, T) {
17   path = numeric(T)
18   path[1] = sample(S, prob = initial_distribution, size = 1)
19   for (n in 1:(T-1)) {
20     path[n+1] = draw_next(path[n])
21   }
22   return(path)
23 }
24 # simulate the entire chain
25 simulate_chain = function(initial_distribution, nsim, T) {
26   initial_state =
27     data.frame(t(replicate(
28       nsim, single_trajectory(initial_distribution, T)
29     )))
30 }
31
32 init_dist = c(1, 0, 0, 0)
33 nsim = 10
34 T = 10
35 df = simulate_chain(init_dist, nsim, T)

```

What do you expect the output of this code to be?

(a) $\text{mean}(\text{df}\$X2)$

$$\begin{aligned}
 \mathbb{E}[X_2 \mid X_1 = 1] &= 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} + 4 \cdot 0 \\
 &= 2.
 \end{aligned}$$

(b) `mean(df$X4==4)`

$$\begin{aligned}
 & \Pr(X_4 = 4 \mid X_1 = 1) \\
 &= \Pr(X_4 = 4, X_3 = 3, X_2 = 3 \mid X_1 = 1) \\
 &+ \Pr(X_4 = 4, X_3 = 3, X_2 = 1 \mid X_1 = 1) \\
 &+ \Pr(X_4 = 4, X_3 = 4, X_2 = 3 \mid X_1 = 1) \\
 &= \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} \cdot 1 \\
 &= \frac{11}{36}.
 \end{aligned}$$

(c) Consider the process at the stopping time

```

1 X_tau=function(x){
2   for (k in 1:length(x)){
3     if (x[k]==2){
4       return(2)
5     }
6     if (x[k]==4){
7       return(4)
8     }
9   }
10  return(3)
11 }
12 Y=apply(df,1,X_tau)

```

What is the expected output of

`mean(Y==4)`

$$\text{Ans: } \frac{1}{2} - \left(\frac{1}{2}\right)^8 + \left(\frac{1}{2}\right) \cdot \left(\frac{1}{3}\right)^8$$

Let $P_k = \Pr(X_k = 4, X_{k-1} \neq 4 \mid X_1 = 1)$, $k = 3, \dots, 10$.

$$\begin{aligned}
 &= \Pr(X_k = 4, X_{k-1} = 3 \mid X_1 = 1) \\
 &= \Pr(X_k = 4 \mid X_{k-1} = 3, X_1 = 1) \cdot \Pr(X_{k-1} = 3 \mid X_1 = 1) \\
 &= \frac{1}{2} \cdot \sum_{i=2}^{k-2} \Pr(X_{k-1} = 3, \dots, X_i = 3, X_{i-1} = 1, \dots, X_2 = 1 \mid X_1 = 1) \\
 &= \frac{1}{2} \cdot \sum_{i=2}^{k-1} \left(\frac{1}{3}\right)^{i-1} \cdot \left(\frac{1}{2}\right)^{k-i-1} = \sum_{i=2}^{k-1} \left(\frac{1}{3}\right)^{i-1} \cdot \left(\frac{1}{2}\right)^{k-i}
 \end{aligned}$$

Page 7 of 9

$$\therefore \Pr\left(\bigcup_{k=3}^{10} \{X_k = 4\}\right) = \sum_{k=3}^{10} P_k = \sum_{i=1}^8 \left[\left(\frac{1}{3}\right)^i \cdot \sum_{k=1}^{9-i} \left(\frac{1}{2}\right)^k \right] = \sum_{i=1}^8 \left(\frac{1}{3}\right)^i - \sum_{i=1}^8 \left(\frac{1}{3}\right)^i \cdot \left(\frac{1}{2}\right)^{9-i} = \frac{1}{2} - \left(\frac{1}{2}\right)^8 + \left(\frac{1}{2}\right) \cdot \left(\frac{1}{3}\right)^8$$

4. A county has 2 large cities and 3 small ones. Any two cities have a direct flight between them, except for the two large ones (since they don't like each other very much). A traveler starts in a large city and moves around randomly by picking one of the available direct flights from their current city at random and taking it. Consider the following R script modelling this:

```

1 # state space
2 S = c(1,2,3,4,5)
3 # transition matrix
4 P <- matrix(c(
5   # From L1
6   0,    0,    1/3,  1/3,  1/3,
7   # From L2
8   0,    0,    1/3,  1/3,  1/3,
9   # From S1
10  1/4,  1/4,  0,    1/4,  1/4,
11  # From S2
12  1/4,  1/4,  1/4,  0,    1/4,
13  # From S3
14  1/4,  1/4,  1/4,  1/4,  0
15 ), nrow = 5, byrow = TRUE)
16
17 # simulate the next position of the chain
18 draw_next = function(s) {
19   sample(S, prob = P[s, ], size = 1)
20 }
21 # simulate a single trajectory of length T from the initial distribution
22 single_trajectory = function(initial_distribution, T) {
23   path = numeric(T)
24   path[1] = sample(S, prob = initial_distribution, size = 1)
25   for (n in 1:(T-1)) {
26     path[n+1] = draw_next(path[n])
27   }
28   return(path)
29 }
30 # simulate the entire chain
31 simulate_chain = function(initial_distribution, nsim, T) {
32   initial_state =
33     data.frame(t(replicate(
34       nsim, single_trajectory(initial_distribution, T))
35     ))
36 }
37
38 init_dist = c(1, 0, 0, 0, 0)
39 nsim = 1000
40 T = 100
41 df = simulate_chain(init_dist, nsim, T)

```

- (a) What is the expected number of flights he or she will take before returning to the initial city for the first time?

```

1 Return_Time = function(x) {
2   for (k in 2:length(x)) {
3     if (x[k] == 1) {

```



```

1     return(k)
2   }
3 }
4   return(length(x))
5 }
6 Time=apply(df,1,Return_Time)
7 print(mean(Time))

```

He/She will have to take
6 flights, thus the computer
will return 7, because $X_7 = 1$.

Let e denote the expected # of flights
to return state 1 if initial state is 1.

Let f denote the expected # of flights
to return state 1 if initial state is 2.

Let g denote the expected # of flights to reach
state 1 if initial state is 3, 4 or 5.

Then

$$\begin{cases} e = 1 + g, \\ g = 1 + \frac{1}{4}f + \frac{1}{2}g \\ f = 1 + g \end{cases} \Rightarrow \begin{cases} e = 6 \\ g = 5 \\ f = 6 \end{cases} \quad \text{Ans: } 6.$$

(b) What is the probability that after 100 flights we are at the opposing city?

```
1 mean(df$X100==2)
```

$$\text{Ans: } \frac{1}{6} - \frac{1}{6} \cdot \left(-\frac{1}{2}\right)^{98}$$

Let p_n denote $P(X_n = 3, 4, 5 | X_1 = 1)$, then

$$p_n = P(X_{n-1} = 1 | X_1 = 1) + P(X_{n-1} = 2 | X_1 = 1) + \frac{1}{2} \cdot p_{n-1}.$$

note that $P(X_{n-1} = 1 | X_1 = 1) = \frac{1}{4} \cdot p_{n-2} = P(X_{n-1} = 2 | X_1 = 1).$

$$\therefore p_n = \frac{1}{2} \cdot p_{n-1} + \frac{1}{2} \cdot p_{n-2}, \quad p_1 = 0, \quad p_2 = 1.$$

Note that $p_n - p_{n-1} = -\frac{1}{2}(p_{n-1} - p_{n-2})$

$$\therefore p_n - p_{n-1} = \left(-\frac{1}{2}\right)^{n-2}$$

$$\text{So } p_n = \sum_{k=2}^n \left(-\frac{1}{2}\right)^{k-2} = \frac{2}{3} - \frac{2}{3} \left(-\frac{1}{2}\right)^{n-1}$$

$$P(X_{100} = 2 | X_1 = 1) = \frac{1}{4} \cdot P(X_{99} = 3, 4, 5 | X_1 = 1) = \frac{1}{6} - \frac{1}{6} \cdot \left(-\frac{1}{2}\right)^{98}.$$