M362M: Introduction to Stochastic Processes

Practice Midterm III

Name: Lee Wei Xuan

UT EID: wl22963

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:

```
# state space
S = c(1,2,3,4,5)
# transition matrix
P <- matrix(c(
  1/2, 0,   0,   0,   1/2,
  0,   1/2, 0,   1/2, 0,
  1/4, 1/4, 1/2, 0,   0,
  0,   1/2, 0,   1/2, 0,
  1/2, 0,   0,   0,   1/2
), nrow = 5, byrow = TRUE)

# simulate the next position of the chain
draw_next = function(s) {
  sample(S, prob = P[s, ], size = 1)
}
# simulate a single trajectory of length T from the initial disribution
single_trajectory = function(initial_distribution,T) {
  path = numeric(T)
  path[1] =sample(S,prob=initial_distribution,size=1)
  for (n in 1:(T-1)) {
    path[n+1] = draw_next(path[n])
  }
  return(path)
}
# simulate the entire chain
simulate_chain = function(initial_distribution,nsim,T) {
  intial_state=
    data.frame(t(replicate(
      nsim, single_trajectory(initial_distribution,T))
    ))
}
```

```
init_dist=c(0,1,0,0,0)
nsim=1000
T=100
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

```
tau_1=function(x){
  for (k in 1:length(x)){
    if (x[k]==1){
     return(k)
    }

  }
  return(length(x))
}
tau_4=function(x){
  for (k in 1:length(x)){
    if (x[k]==4){
     return(k)
    }
  }
  return(length(x))
}
```

```
T_1=apply(df,1,tau_1)
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 $k=1, \cdots, 100$ if we start with $X_1 = 2$.

Thus we expect to see 100.

(c) `mean(T_4)`

Let ν be the stopping time

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

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

$$= 1 \cdot 0 + \sum_{k=2}^{99} k \cdot \left(\tfrac{1}{2}\right)^{k-1} + 100 \cdot \Big[\underbrace{\left(\tfrac{1}{2}\right)^{99}}_{X_1 = \cdots = X_{100} = 2} + \underbrace{\left(\tfrac{1}{2}\right)^{99}}_{X_1 = \cdots = X_{99} = 2,\ X_{100} = 4}\Big]$$

$$= 0 + \frac{1}{1-\frac{1}{2}} + \frac{\left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^{99}}{\left(1-\frac{1}{2}\right)^2} - \frac{99 \cdot \left(\frac{1}{2}\right)^{99}}{1-\frac{1}{2}} + 100 \cdot \left(\tfrac{1}{2}\right)^{98}$$

$$= 3 - \left(\tfrac{1}{2}\right)^{97} + \left(\tfrac{1}{2}\right)^{98}$$

2. Consider the following R script:

```
# state space
S = c(1,2,3)
# transition matrix
P <- matrix(c(
  1/2, 1/2, 0,
  1/5, 2/5, 2/5,
  1/5, 1/5, 3/5
  ), nrow = 3, byrow = TRUE)

# simulate the next position of the chain
draw_next = function(s) {
  sample(S, prob = P[s, ], size = 1)
}
# simulate a single trajectory of length T from the initial disribution
single_trajectory = function(initial_distribution,T) {
  path = numeric(T)
  path[1] =sample(S,prob=initial_distribution,size=1)
  for (n in 1:(T-1)) {
    path[n+1] = draw_next(path[n])
  }
  return(path)
}
# simulate the entire chain
simulate_chain = function(initial_distribution,nsim,T) {
  intial_state=
    data.frame(t(replicate(
      nsim, single_trajectory(initial_distribution,T))
    ))
}
init_dist=c(1,0,0)
nsim=1000
T=1000
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)`

$P(X_2 = 1 \mid X_1 = 1) = 1/2.$

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

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

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

So

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

$$\therefore \ \underset{=0.2857}{\pi_1^* = \frac{2}{7}}, \ \underset{=0.3571}{\pi_2^* = \frac{5}{14}}, \ \underset{=0.3571}{\pi_3^* = \frac{5}{14}}$$

(c) `mean(df$X1000)`

Suppose X_{1000} follow the distribution π_*

Then

$$E[X_{1000}] = 1\cdot\frac{2}{7} + 2\cdot\frac{5}{14} + 3\cdot\frac{5}{14}$$

$$= \frac{29}{14}.$$

$$= 2.0714.$$

3. Consider the following R script:

```
# state space
S = c(1,2,3,4)
# transition matrix
P <- matrix(c(
  1/3, 1/3, 1/3, 0,    # Transitions from state 1
  0,   1,   0,   0,    # Transitions from state 2
  0,   0,   1/2, 1/2,  # Transitions from state 3
  0,   0,   0,   1     # Transitions from state 4
), nrow = 4, byrow = TRUE)

# simulate the next position of the chain
draw_next = function(s) {
  sample(S, prob = P[s, ], size = 1)
}
# simulate a single trajectory of length T from the initial disribution
single_trajectory = function(initial_distribution,T) {
  path = numeric(T)
  path[1] =sample(S,prob=initial_distribution,size=1)
  for (n in 1:(T-1)) {
    path[n+1] = draw_next(path[n])
  }
  return(path)
}
# simulate the entire chain
simulate_chain = function(initial_distribution,nsim,T) {
  intial_state=
    data.frame(t(replicate(
      nsim, single_trajectory(initial_distribution,T))
    ))
}

init_dist=c(1,0,0,0)
nsim=10
T=10
df = simulate_chain(init_dist,nsim,T)
```

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

(a) `mean(df$X2)`

$$\mathbb{E}[X_2 \mid X_1 = 1] = 1 \cdot \tfrac{1}{3} + 2 \cdot \tfrac{1}{3} + 3 \cdot \tfrac{1}{3} + 4 \cdot 0$$

$$= 2.$$

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

$$\Pr(X_4 = 4 \mid X_1 = 1)$$

$$= \Pr(X_4=4, X_3=3, X_2=3 \mid X_1=1)$$

$$+ \Pr(X_4=4, X_3=3, X_2=1 \mid X_1=1)$$

$$+ \Pr(X_4=4, X_3=4, X_2=3 \mid X_1=1)$$

$$= \tfrac{1}{3}\cdot\tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{3}\cdot\tfrac{1}{3}\cdot\tfrac{1}{2} + \tfrac{1}{3}\cdot\tfrac{1}{2}\cdot 1$$

$$= \tfrac{11}{36}.$$

(c) Consider the process at the stopping time

```
X_tau=function(x){
  for (k in 1:length(x)){
    if (x[k]==2){
      return(2)
    }
    if (x[k]==4){
      return(4)
    }
  }
  return(3)
}
Y=apply(df,1,X_tau)
```

What is the expected output of

```
mean(Y==4)
```

ans : $\tfrac{1}{2} - \left(\tfrac{1}{2}\right)^8 + \left(\tfrac{1}{2}\right)\cdot\left(\tfrac{1}{3}\right)^8$

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

$$= \Pr(X_k = 4, X_{k-1} = 3 \mid X_1 = 1)$$

$$= \Pr(X_k = 4 \mid X_{k-1} = 3, X_1 = 1) \cdot \Pr(X_{k-1} = 3 \mid X_1 = 1)$$

$$= \tfrac{1}{2} \cdot \sum_{i=2}^{k-2} \Pr(X_{k-1} = 3, \dots, X_i = 3, X_{i-1} = 1, \dots, X_2 = 1 \mid X_1 = 1)$$

$$= \tfrac{1}{2} \cdot \sum_{i=2}^{k-1} \left(\tfrac{1}{3}\right)^{i-1} \cdot \left(\tfrac{1}{2}\right)^{k-i-1} = \sum_{i=2}^{k-1} \left(\tfrac{1}{3}\right)^{i-1} \cdot \left(\tfrac{1}{2}\right)^{k-i}$$

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

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:

```
# state space
S = c(1,2,3,4,5)
# transition matrix
P <- matrix(c(
  # From L1
  0,    0,    1/3,  1/3,  1/3,
  # From L2
  0,    0,    1/3,  1/3,  1/3,
  # From S1
  1/4,  1/4,  0,    1/4,  1/4,
  # From S2
  1/4,  1/4,  1/4,  0,    1/4,
  # From S3
  1/4,  1/4,  1/4,  1/4,  0
), nrow = 5, byrow = TRUE)

# simulate the next position of the chain
draw_next = function(s) {
  sample(S, prob = P[s, ], size = 1)
}
# simulate a single trajectory of length T from the initial disribution
single_trajectory = function(initial_distribution,T) {
  path = numeric(T)
  path[1] =sample(S,prob=initial_distribution,size=1)
  for (n in 1:(T-1)) {
    path[n+1] = draw_next(path[n])
  }
  return(path)
}
# simulate the entire chain
simulate_chain = function(initial_distribution,nsim,T) {
  intial_state=
    data.frame(t(replicate(
      nsim, single_trajectory(initial_distribution,T))
    ))
}

init_dist=c(1,0,0,0,0)
nsim=1000
T=100
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?

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

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

```
        return(k)
      }
    }
    return(length(x))
}
Time=apply(df,1,Return_Time)
print(mean(Time))
```

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?

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

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

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

$$P_n = P(X_{n-1}=1 \mid X_1 = 1) + P(X_{n-1} = 2 \mid X_1 = 1) + \frac{1}{2}\cdot P_{n-1}.$$

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

$$\therefore\ P_n = \frac{1}{2}\cdot P_{n-1} + \frac{1}{2}\cdot P_{n-2},\quad P_1 = 0,\ P_2 = 1.$$

Note that $P_n - P_{n-1} = -\frac{1}{2}(P_{n-1} - P_{n-2})$

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

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

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