Simulating Discrete Markov Chains

University of Toronto, ACT 350

Prof. S. Pesenti

27/10/2020

An introductory example

Assume a student has lectures every day at the University of Toronto and is either punctual or late. The student is late with probability 0.3, when he was late the day before, and late with probability 0.4, when he was punctual the day before. Being late or punctual two days before does not influence his punctuality.

1. Defining the Markov Chain

- State space: $S = \{l, p\}$, l means late, p means punctual
- $\{X_n,\ n = 0, 1, \dots\}$, where the event $\{X_n = l\}$ means that the student is late at day n.
- Fulfills the Markov property since the event that the student is late or punctual only depends whether the student was late or punctual the day before. It holds for $n \geq 0$

$$
\begin{aligned}
p_{l,l} &= P(X_{n+1} = l | X_n = l) = 0.3 \\
p_{p,l} &= P(X_{n+1} = l | X_n = p) = 0.4 \\
p_{l,p} &= P(X_{n+1} = p | X_n = l) = 1 - P(X_{n+1} = l | X_n = l) = 0.7 \\
p_{p,p} &= P(X_{n+1} = p | X_n = p) = 1 - P(X_{n+1} = l | X_n = p) = 0.6
\end{aligned}
$$

- transition matrix $P = \begin{pmatrix} p_{l,l} & p_{l,p} \\ p_{p,l} & p_{p,p} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix}$.

We call $n = 0$ the induction day and $n > 0$ represents the n-th day of lectures.

Numerical implementation: R code

```r
# initialise the transition matrix
P <-  matrix(c(0.3, 0.4, 0.7, 0.6), nrow = 2)
rownames(P) <- c("l","p")
colnames(P) <- c("l","p")
P
```

```
##     l   p
## l 0.3 0.7
## p 0.4 0.6
```

```
# check that all rows sum up to 1
rowSums(P)
```

```
## l p
## 1 1
```

2. Transition probabilities and stationary distribution

Next we calculate the conditional probabilities for a student being late at a later day in the semester given that we know whether the student was late on induction day.

Question: What are the conditional transition probabilities that a student is late or punctual on the second day? That is e.g. $p^2_{l,l} = P(X_1 = l|X_0 = l)$, $p^2_{l,p}, p^2_{p,l}$ and $p^2_{p,p}$.

Question: How about the probabilities for the 10^{th} day?

Question: How about the probabilities for the 50^{th} day?

For this, we have to calculate the transition probabilities for day 2 ($P * P$), and for day 10 (P^{10}) and for day 50 (P^{50}). For this we can use the Chapman-Kolmogorov equation to calculate these transition probabilities. If we denote by $P^1 = (p_{i,j})_{i,j,\in S}$ the transition probability matrix for the first day. Then, $P^2 = P^1 * P^1$ by Chapman-Kolmogorov and $P^n = (P^1)^n$, that is the matrix P^1 multiplied n times with itself.

```
# transition probabilities on 2 days
P %*% P
```

```
##      l    p
## l 0.37 0.63
## p 0.36 0.64
```

```
# check that all rows sum up to 1
rowSums(P)
```

```
## l p
## 1 1
```

Question: What is the interpretation of `(P %*% P)[1,1]` = 0.37?

```
# transition probabilities on 10 days
P %^% 10
```

```
##           l         p
## l 0.3636364 0.6363636
## p 0.3636364 0.6363636
```

```
# transition probabilities on 50 days
P %^% 50
```

```
##           l         p
## l 0.3636364 0.6363636
## p 0.3636364 0.6363636
```

Question: What do you observe?

We see numerically that the rows of the transition matrices always add up to 1. Further, we observe that the transition matrix for 10 days and for 50 days are numerically equal. Let $t = 0$ correspond to the induction day and $t = 1$ to the first day of lectures. Then we have that, if the student was late at the induction day, he will be late at the end of the semester with probability 0.3636364, independent of whether he was late or not during the term. Recall that $p_{ll}^{50} = P(X_{50} = l|X_0 = l)$ is the probability that the student is late on the 50th day, given he was late at the induction day. Numerically, we observe that $p_{ll}^{10} = p_{ll}^{11} = \cdots = p_{ll}^{50} = \cdots p_{ll}^{n} = \cdots$ for all n large. Thus, the probability of being late on a day later in the semester given that the student was late on the first day, is independent whether the student was late at during the semester and only depends on whether the student was late on induction day or not.

Question: Does this mean that (0.3636364, 0.6363636) is a stationary distribution for the Markov Chain?

Question: What properties of the Markov Chain imply that the Markov Chain has a *stationary distribution*?

Question: How many stationary distributions does the Markov Chain have?

Question: Calculate analytically the stationary distribution.

```
# the analytical stationary distribtuion is
pi <- c(4/11, 7/11)
names(pi) <- c("l", "p")
# we veryfy that pi is a stationary distribtuion, that is pi * P = pi
t(pi) %*% P
```

```
##              l         p
## [1,] 0.3636364 0.6363636
```

```
pi
```

```
##         l         p
## 0.3636364 0.6363636
```

```
all.equal(t(pi) %*% P, t(pi))
```

```
## [1] TRUE
```

3. Initial distribution

Let us look at students whose transition probabilities are described by P but have different probabilities to be punctual at induction day ($t = 0$).

- *student 1*: probabilities at induction day $(0.5, 0.5)$ (random student)
- *student 2*: probabilities at induction day $(0.2, 0.8)$ (good student)
- *student 3*: probabilities at induction day $(0.7, 0.3)$ (bad student)

student 1

```
# probability of student 1 at induction day
student1_x0 <- c(0.5, 0.5)

# probability at the first day
student1_x0 %*% P
```

```
##         l    p
## [1,] 0.35 0.65
```

```
# probability at the 5th day
student1_x0 %*% (P %^% 5)
```

```
##             l        p
## [1,] 0.363635 0.636365
```

```
# probability at the 100th day
student1_x0 %*% (P %^% 100)
```

```
##              l         p
## [1,] 0.3636364 0.6363636
```

Observe that *student 1* has probability 0.5 to be punctual at induction day. On the first day of lectures however, *student 1* has probability 0.65 to be punctual. *Student 1* is punctual on the 100-th day with probability 0.6363636.

student 2
Now let us look at the "good" *student 2* who started with probability of 0.8 to be punctual at induction day.

```
# probabilities of student 2 at induction day
student2_x0 <- c(0.2, 0.8)

# probabilities at the first day
student2_x0 %*% P
```

```
##         l    p
## [1,] 0.38 0.62
```

```
# probabilities at the 5th day
student2_x0 %*% (P %^% 5)
```

```
##             l        p
## [1,] 0.363638 0.636362
```

```
# probabilities at the 100th day
student2_x0 %*% (P %^% 100)
```

```
##              l         p
## [1,] 0.3636364 0.6363636
```

We observe, that even though the "good" *student 2* started with a higher probability of being punctual on induction day, the probability that the student is late on the 100-th day is equal to 0.6363636, which is equal the corresponding probability of *student 1*.

student 3
How about *student 3* who only has a probability of 0.3 to turn up punctual at induction day? Will the student improve over the semester?

```
# probabilities of student 3 at induction day
student3_x0 <- c(0.7, 0.3)

# probabilities at the first day
student3_x0 %*% P
```

```
##         l    p
## [1,] 0.33 0.67
```

```
# probabilities at the 5th day
student3_x0 %*% (P %^% 5)
```

```
##             l        p
## [1,] 0.363633 0.636367
```

```
# probabilities at the 100th day
student3_x0 %*% (P %^% 100)
```

```
##              l         p
## [1,] 0.3636364 0.6363636
```

We observe that at the end of the semester, *student 1* (random student), *student 2* (good student) and *student 3* (bad student) have the same probability (0.6363636) to come punctually to the lectures.

Question: Does this hold for any student, whose transition probabilities are described by P? That is, does the initial distribution of the student not influence the limiting distribution?

Numerically, and also analytically (**Why?**), it holds that

$$\lim_{n\to\infty}(\nu, 1-\nu) * P^n = (\nu, 1-\nu) * \begin{pmatrix} 0.3636364 & 0.6363636 \\ 0.3636364 & 0.6363636 \end{pmatrix} = (0.3636364, 0.6363636), \tag{1}$$

for any initial distribution $\nu, 1-\nu,\ \nu \in [0,1]$ on the state space $S = \{l, p\}$.

4. Sample path

So far we only looked how the probabilities of being punctual or late changes from day to day or over the semester. In this section we look at sample paths of the Markov Chain. If we ask a student to note down every single day whether he was late or punctual, he will provide us with an (infinite) sequence consisting of l's and p's – his sample path.

Simulating a sample path:

1. Set an initial value for x_0. (Is the student punctual or late at induction day?) To be able to use the matrix notation, the initial value needs to be a vector. For example $x_0 = (1, 0)$ means the student is late at induction day.

2. for $n \geq 1$ do:

   a) calculate $x_{n-1} * P$ (the distribution of the student to be late/punctual on the n^{th} day given the outcome of the previous day)

b) generate a realisation from a random variable with distribution $x_{n-1} * P$

c) set the outcome of the random variable equal to x_n

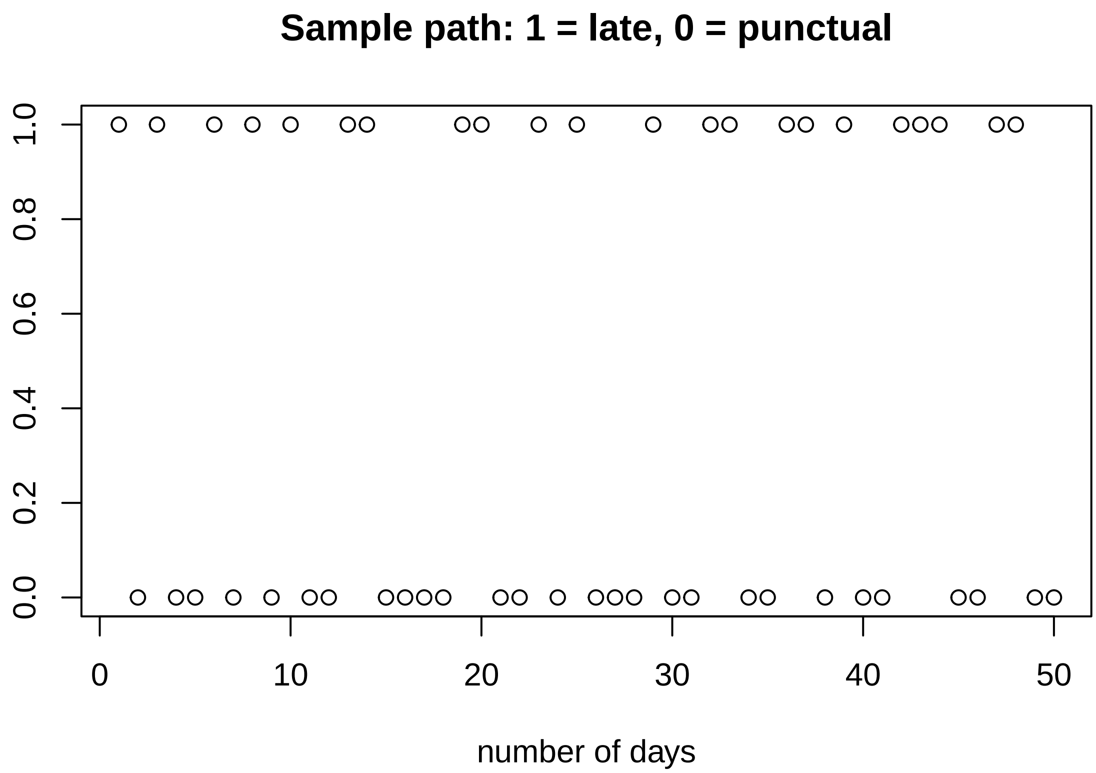
The collection $x_0, x_1, \dots$ is the sample path of the student.

A simulated sample path: R code
For simplicity, denote $l = 1$ and $p = 0$. Choose $x_0 = (1, 0)$, that is the student is late on the first day.

```
# set the seed for generating random variables (for reproducibility)
# set.seed(2020)
# number of days (length of simulated sample path)
n <- 50
# initialise the sample path - 1 in sample_path mean late, 0 means punctual.
sample_path <- rep(0, n)
prob_late <- rep(0, n)
# initial value at induction day
x <- c(1, 0)
# probability that the student is late on the next day
prob <- x * P
# generate a realisation with distribution P
sample_path[1] <- rbinom(1, 1, P[1,1])
# determine the vector of whether the student is late or not.
x <- c(sample_path[1], if(sample_path[1] == 0){1} else{0})

for(i in 2:n){
  prob <- (x %*% P)
  # generate a Bernoulli random variable with probability = late
  sample_path[i] <- rbinom(1, 1, prob)
  x <- c(sample_path[i], if(sample_path[i] == 0){1} else{0})
}

plot(seq(1, n), sample_path,
     xlab = "number of days", ylab = " ", type= "p",
     main = "Sample path: 1 = late, 0 = punctual")
```



The plot shows the generated sample path, that displays whether the student is late (1) or punctual (0) for the next 50 day.

Question If you rerun the above code to simulate a sample path, you should always get something different. Why?

Question Rerun the above code to simulate a sample path changing n, the length of the sample path, and/or the initial value.

Question You can also rerun the whole document using different transition probabilities. (However, the text might be sometimes inconsistent…)