Stochastic Processes

Kevin Chang

March 31, 2022

1

(a). Suppose that coin 1 has probability 0.7 of coming up heads, and coin 2 has probability 0.6 of coming up heads. If the coin flipped today comes up heads, then we select coin 1 to flip tomorrow, and if it comes up tail, then we select coin 2 to flip tomorrow. If the coin initially flipped is equally likely to be coin 1 or coin 2, then what is the probability that the coin flipped the on the third day after the initial flip coin is 1.
(b). Suppose that the coin flipped on Monday comes up heads. What is the probability that the coin flipped on Friday of the same week also comes up heads?

- $P = \begin{bmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{bmatrix}$
- (a) $[0.5\ 0.5]P^3 = [0.6665\ 0.3335]$
- (b) $[1\ 0]P^3 = [0.6667\ 0.3333]$

2

A professor continually gives exam to his students. He can give three possible types of exams, and her class is graded as either having done well or badly. Let p_i denote the probability that the class done well on type i exam, and suppose that $p_1 = 0.3, p_2 = 0.6$, and $p_3 = 0.9$. If the class does well on an exam, then the next exam is equally likely to be any of the three types. If the class does badly, then the next exam is always type 1. What proportion of exams are type $i, i = 1, 2, 3$?

- $P = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.6 & 0.2 & 0.2 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}$
- stationary distribution: p
  - p is eigenvector corresponding to eigenvalue 1 of P^T
  - $p = [\frac{5}{7}\ \frac{1}{7}\ \frac{1}{7}]$

3

A DNA nucleotide has any of the 4 values. A standard model for a mutational change of the nucleotide at a spcific location is a Markov chain model that supposes that in going from period to period, the nucleotide does not change with probability $1 - 3\alpha$ and if it does change then it is equally likely to change to any of the other 3 values, for some $0 < \alpha < \frac{1}{3}$

(a) Show that $P^n_{1,1} = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n$.

(b) What is the long run proportion of time the chain is in each state?

- $P = \begin{bmatrix} 1-3\alpha & \alpha & \alpha & \alpha \\ \alpha & 1-3\alpha & \alpha & \alpha \\ \alpha & \alpha & 1-3\alpha & \alpha \\ \alpha & \alpha & \alpha & 1-3\alpha \end{bmatrix}$
- (a)
  - $P^n = \frac{1}{4}\begin{bmatrix} 3(1-4\alpha)^n+1 & 1-(1-4\alpha)^n & 1-(1-4\alpha)^n & 1-(1-4\alpha)^n \\ 1-(1-4\alpha)^n & 3(1-4\alpha)^n+1 & 1-(1-4\alpha)^n & 1-(1-4\alpha)^n \\ 1-(1-4\alpha)^n & 1-(1-4\alpha)^n & 3(1-4\alpha)^n+1 & 1-(1-4\alpha)^n \\ 1-(1-4\alpha)^n & 1-(1-4\alpha)^n & 1-(1-4\alpha)^n & 3(1-4\alpha)^n+1 \end{bmatrix}$

    (by Wolfram Alpha)
  - $P^n_{1,1} = \frac{1}{4} + \frac{3}{4}(1 - 4\alpha)^n$
- (b) it is equally distributed: stationary distribution is $\frac{1}{4}[1, 1, 1, 1]$

4

› Let $P^n_{ij} = P_i[X_n = j] = P[X_n = j | X_0 = i]$.

Let $f^n_{ij} = P[X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j | X_0 = i]$
ie., the probability that the first transition into occurs at time n. $f^0_{ij} = 0$.

Show that

$$P^n_{ij} = \sum_{k=0}^{n} f^k_{ij} P^{n-k}_{jj}.$$

- $P^n_{ij} = P[X_n = j | X_0 = i] = P[X_n = j, X_1 = j | X_0 = i] + P[X_n = j, X_1 \neq j | X_0 = i]$
  $= P[X_n = j | X_1 = j, X_0 = i] P[X_1 = j | X_0 = i] + P[X_n = j, X_1 \neq j | X_0 = i]$
  $= P^{n-1}_{jj} f^k_{ij} + P[X_n = j, X_2 = j, X_1 \neq j | X_0 = i] + P[X_n = j, X_2 \neq j, X_1 \neq j | X_0 = i]$
  $= P^{n-1}_{jj} f^2_{ij} + P^{n-2}_{jj} f^2_{ij} + P[X_n = j, X_2 \neq j, X_1 \neq j | X_0 = i]$
  $\ldots$
  $= \sum_{k=0}^{n} P^{n-k}_{jj} f^k_{ij}$

5

Jobs arrive at a processing center in accordance with a Poisson process with rate λ. However, the center has waiting space for only N jobs and so an arriving job finding N others waiting goes away. At most 1 job per day can be processed, and the processing of this job must start at the beginning of the day. Thus, if there are any jobs waiting for processing at the beginning of a day, then one of them is processed that day, and if no jobs are waiting at the beginning of a day then no jobs are processed that day. Let X_n denote the number of jobs at the center at the beginning of day n.

(a) Find the transition probabilities of the Markov chain $\{X_n, n \geq 0\}$.

(b) Is this chain ergodic? Explain.

(c) Write the equations for the stationary probabilities.

- $Y \sim \text{Poisson}(\lambda)$
- (a) $P = \begin{bmatrix} P[Y=0] & P[Y=1] & \dots & \dots & P[Y=N] \\ P[Y=0] & P[Y=1] & \dots & \dots & P[Y=N] \\ 0 & P[Y=0] & P[Y=1] & \dots & P[Y \geq N-1] \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & P[Y=1] & P[Y \geq 1] \end{bmatrix}$
- (b) yes it is ergodic, since it is aperiodic and irreducible

6

Consider a population of size n, some of whom are infected with a certain virus. Suppose that in an interval of length h any specified pair of individuals will independently interact with probability $\lambda h + o(h)$. If exactly one of the individuals involved in the interaction is infected then the other one becomes infected with probability α. If there is a single individual infected at time 0, find the expected time at which the entire population is infected.

$h = \delta t$

- $R = \begin{bmatrix} -\lambda & \lambda(1-\alpha) & \lambda\alpha & \dots & \dots & 0 \\ 0 & -\lambda & \lambda(1-\alpha) & \lambda\alpha & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix}$
- Suppose
  - N is the number of infection
  - X_i is the number of infected perseon in each time
- $\sum_{i=1}^{N} X_i - N\mathbb{E}[X_i]$ is a martingale
- $\mathbb{E}[N] = \frac{n-1}{\mathbb{E}[X_i]} = \frac{n-1}{1+a}$
- Expected time at which the whole population is infected: $\lambda \frac{n-1}{1+a}$

7

Consider a Markov chain $\{X_n, n \geq 0\}$ with $P_{NN} = 0$. Let $P(i)$ denote the probability that this chain eventually enters state N given that it starts in state i. Show that $\{P(X_n), n \geq 0\}$ is a martingale.

- A is a random variale such that
  $A = \begin{cases} 1 & \text{if this chain eventually enters state } N \\ 0 & \text{otherwise} \end{cases}$
- $P(X_n) = \mathbb{E}[A|X_n]$
- $\mathbb{E}[P(X_{n+1})|X_n, \ldots, X_1] = \mathbb{E}[\mathbb{E}[A|X_{n+1}]|X_n, \ldots, X_1] = \mathbb{E}[A|X_n] = P(X_n)$
- $\mathbb{E}[P(X_n)] \leq 1$

8

Let $X(n)$ denote the size of the nth generation of a branching process, and let π denote the probability that such a process, starting with a single individual, eventually goes extinct. Show that $\{\pi^{X(n)}, n \geq 0\}$ is a martingale.

- A is a random variale such that
  $A = \begin{cases} 1 & \text{if this chain eventually enters state } N \\ 0 & \text{otherwise} \end{cases}$
- $\pi^{X(n)} = \mathbb{E}[\pi|X(n), \ldots, X(0)]$
- $\mathbb{E}[\pi^{X(n+1)}|X(n), \ldots, X(0)] = \mathbb{E}[\mathbb{E}[\pi|X(n+1), X(n), \ldots, X(0)]|X(n), \ldots, X(0)]$
  $= \mathbb{E}[\pi|X(n), \ldots, X(0)] = \pi^{X(n)}$
- $\mathbb{E}[\pi^{X(n)}|X(n-1), \ldots, X(0)] \leq 1$

9

Let $Z_n = \prod_{i=1}^{n} X_i$, where X_i, $i \geq 1$ are independent random variables with

$$P\{X_i = 2\} = P\{X_i = 0\} = 1/2.$$

Let $N = \text{Min}\{n\colon Z_n = 0\}$. Is the martingale stopping theorem applicable? If so, what would you conclude? If not, why not?

- No
  - By first rule
    - We can not find a constant k such that $P[N < k] = 1$
  - By second rule
    - We can not find a constant k such that $P[\prod_{i=1}^{n} X_i < k] = 1$

- By third rule
  * $\mathbb{E}[N] = \sum_{i=1}^{\infty} i 2^{-i} = 2$
  * We can not find a constant k such that $\mathbb{E}[|Z_{n+1} - Z_n||Z_n, \ldots, Z_1] = Z_n < k$

10

Let X denote the number of successes in n independent Bernoulli trials, with trial i resulting in a success with probability p_i. Give an upper bound for $P\left\{\left|X - \sum_{i=1}^{n} p_i\right| \geq a\right\}$.

- S_i i-th Bernoulli trial
- $Z_n = \sum_{i=1}^{n} S_i - \sum_{i=1}^{n} p_i$ is a martingale
  - $\mathbb{E}[Z_n] < n$
  - $\mathbb{E}[Z_{n+1}|S_n, \ldots, S_1] = \mathbb{E}[Z_n + S_{n+1} - p_{n+1}|S_n, \ldots, S_1] = Z_n$
- By Azuma's inequality: $P[|Z_n| \geq a] \leq 2e^{-\frac{a^2}{2n}}$
  - $-1 \leq Z_{i+1} - Z_i \leq 1$

11

6.11. Consider a gambler who at each gamble is equally likely to either win or lose 1 unit. Suppose the gambler will quit playing when his winnings are either A or $-B$, $A > 0$, $B > 0$.

Let

$$N = \min\{n: \quad S_n \geq A \text{ or } S_n \leq -B\}.$$

Show that $E[N] < \infty$. (*Hint:* Argue that there exists a value k such that $P\{S_k > A + B\} > 0$. Then show that $E[N] \leq kE[G]$, where G is an appropriately defined geometric random variable.)

Use an appropriate martingale to show that the expected number of bets is AB.

- E is an event that the gambler wins successively $A + B$ times
- $P[E] = 2^{A+B}$
- $\mathbb{E}[\#\text{ of trials until } E] = \frac{1}{2^{A+B}}$
- $\mathbb{E}[N] \leq \frac{A+B}{2^{A+B}}$
- $Z_n = (\sum_{i=1}^{n} S_i)^2 - n$ is a martingale
  - $\mathbb{E}[|Z_n|] < n^2$
  - $\mathbb{E}[Z_{n+1}|S_n, \ldots, S_1] = \mathbb{E}[Z_n + S_{n+1}^2 + 2S_{n+1}(\sum_{i=1}^{n} S_i) - 1|S_n, \ldots, S_1] = Z_n$
- By third rule

- $\mathbb{E}[N] \leq \infty$
- $\mathbb{E}[|Z_{n+1} - Z_n||Z_n, \ldots, Z_1]$
  $= \mathbb{E}[|S_{n+1}^2 + 2S_{n+1}(\sum_{i=1}^{n} S_i) - 1||S_n, \ldots, S_1] \leq 1 + 2(A+B) + 1$

- $\mathbb{E}[Z_N] = 0$
  - $\mathbb{E}[Z_N] = \mathbb{E}[(\sum_{i=1}^{N} S_i)^2] - \mathbb{E}[N] = 0$
  - $\mathbb{E}[N] = \frac{A}{A+B}B^2 + \frac{B}{A+B}A^2 = AB$

12

6.17. Suppose that 100 balls are to be randomly distributed among 20 urns. Let X denote the number of urns that contain at least five balls. Derive an upper bound for $P\{X \geq 15\}$. (Your upper bound should be a number <1, like 0.6.)
Using a computer, simulate at least 10000 trials of this process to estimate $\mathbb{P}[X \geq 15]$. How good is your bound?

- $\mathbb{E}[X]$
  - $\mathbb{E}[X] = 20 \times (1 - \frac{19^{100} + \binom{100}{1}19^{99} + \binom{100}{2}19^{98} + \binom{100}{3}19^{97} + \binom{100}{4}19^{96}}{20^{100}}) \approx 11.2803)$
- B_i is the result for i-th ball
- By Azuma's inequality: $P[|Z_n - \mathbb{E}[X]| \geq a] \leq 2e^{-\frac{a^2}{2n}}$
  - $P[|Z_n - \mathbb{E}[X]| \geq a] \leq 2e^{-\frac{a^2}{2n}}$
  - $P[X \geq 15] \leq 2e^{-\frac{(15-\mathbb{E}[X])^2}{200}} \approx 0.93316$
- Simulation

```python
import numpy as np
import matplotlib
import matplotlib.pyplot as plt

m = 100
n = 20
def expr():
    test = np.zeros(n)
    res = np.random.randint(n, size=m)
    for i in range(m):
        test[res[i]] += 1
    return np.sum(test>=15) > 0

trial = 10000
c = 0
for t in range(trial):
    res = expr()
    if res:
        c += 1

print(c / trial)
```

- Result

```
➜ python3 simulation.py
0.0022
```