Section 5 (Stat 171)

- TF: Jiaze Qiu
  - Email: jiazeqiug.harvard.edu
  - Section/OH: Wed 9-12 pm (ET)
- TF: Fernando Vicente
  - Email: fernando_vicente@fas.harvard.edu
  - Section/OH: Mon 7-9 pm and Thu 8-10 pm (ET)

- Acknolegement: This handout is partially based on notes created by Lisa Ruan and Christy Huo.

- All the section materials (handouts & solutions) can be found either on Canvas or here.

1 Topic list for midterm 1

- Basic probability (see Section 1)
- Definition of a Markov chain & Markov property
- Definition of a transition matrix
- n-step transition matrix & Chapman-Kolmogorov relationship
- Limiting distribution & stationary distribution
- First-step analysis
- Limit theorem for regular Markov chains
- Simple random walk (SRW)
- Random walk on graphs
- Definition of accessibility & communication relation
- Definition of irreducibility
- Recurrence & transience (class properties)
- Limit theorem for finite irreducible Markov chains
- Periodicity (class property)
- Fundamental limit theorem for ergodic Markov chains
- Time reversibility
- Absorbing chains
- Definition of branching processes
- Mean generation size and variance of generation size for three cases
- Probability generating functions and properties
- Extinction probability theorem

2 Practice problems

2.1 Recurrence

Suppose a particle moves along the integers starting at 0. During each unit of time, it moves up 2 (i.e. +2) with probability 1/3 or down 1 (i.e. -1) with probability 2/3.
(a) Find the period of each state.
(b) Is 0 a recurrent state?
Solutions.
(b) Recall that we only need to show that $\sum_{k=0}^{\infty} P_{00}^k = \infty$ to prove 0 is recurrent. We have

$$\sum_{k=0}^{\infty} P_{00}^k = \sum_{k=0}^{\infty} P_{00}^{3k} = \sum_{k=0}^{\infty} \frac{(3k)!}{k!(2k)!}(\frac{1}{3})^k(\frac{2}{3})^{2k}$$

Note that

$$n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$$

Thus

$$P_{00}^{3k} \sim C\frac{1}{\sqrt{k}}$$

which implies

$$\sum_{k=0}^{\infty} P_{00}^k = \sum_{k=0}^{\infty} P_{00}^{3k} = \infty$$

2.2 First-step analysis

A particle moves on a circle through seven points, marked as 0,1,2,3,4,5 and 6 in clockwise order. At each step, it has probability 1/2 of moving clockwise and probability 1/2 of moving counterclockwise. Let X_n denote the location on the circle after the n-th step.
(a) What is the expected number of steps the particle takes to return to the starting position?
(b) Starting in state 0, what is the probability that all other positions are visited before the particle returns to its starting position?
Solutions. (Outline)
(a) Let $T_j = \inf\{t > 0 : X_t = 0\}$. Define

$$x_i = E[T_j|X_0 = j]$$

Note that $x_i = x_{7-i}$. By first-step analysis, we have the following equations

$$\begin{cases} x_1 = 0.5 + 0.5(1 + x_2) \\ x_2 = 0.5(1 + x_1) + 0.5(1 + x_3) \\ x_3 = 0.5(1 + x_2) + 0.5(1 + x_3) \end{cases}$$

Thus

$$x_0 = 1 + x_1 = 7$$

(b) Same logic - but note that we no longer have "$x_i = x_{7-i}$" type of symmetry.

2.3 Branching process

Consider a branching process where an individual gives rise to a random number of offspring according to the geometric distribution on $\{0, 1, 2, ...\}$, namely there are j offspring with probability $p(1-p)^j$ for $j = 0, 1, 2, ...$ Suppose the population starts with one individual.
(a) What condition on p ensures the population will die out for sure?
(b) For other values of p, what is the probability of extinction?
Solutions.
(a) It suffices to consider epectation of the offspring distribution:

$$E_{X\sim \text{Geometric}(p)} X = \frac{1-p}{p}$$

(b) PGF of the offspring distribution:

$$G(s) = Es^X = \sum_{k=0}^{\infty} p(1-p)^k s^k = p\sum_{k=0}^{\infty}[s(1-p)]^k = \frac{p}{1-s(1-p)}$$

Hint: consider the smallest root of $G(s) = s$

2.4 Application of Markov Chain theory

A clinical trial is run to determine which of two drugs is more effective in treating a particular disease. Independently, as pairs of patients enter the trial, one randomly gets drug A and the other gets drug B. The outcome is (A_i, B_i) where $A_i = 1$ if drug A cured the patient taking drug $A, 0$ if not. $B_i = 1$ if drug B cured the other patient. Let

$$S_n = \sum_{i=1}^{n} (A_i - B_i)$$

denote the net difference in the number of cures out of n pairs, so that $S_n > 0$ means drug A is doing better than B. Suppose the trial is run until the first trial n such that $|S_n| = M$, where M is some fixed number. If M is large, this seems to be a good way of deciding which drug is better. But it is still possible to make an error. Let $a = P\{A_i = 1\}$ and $b = P\{B_i = 1\}$ denote the unknown cure probabilities, and assume the A_i and B_i form independent sequences, each being an i.i.d. Bernoulli sequence with success probabilities a and b, respectively.

(a) Let

$$p = P\{A_i - B_i = 1 \mid A_i - B_i \neq 0\}$$

In terms of a and b, what is p?

(b) Assuming $a > b$, let α_M be the chance the trial stops by declaring B to be better, the wrong conclusion. What is α_M in terms of p and M? What does it tell you?

Solutions.

(a)

$$p = \frac{P(A_i - B_i = 1, A_i - B_i \neq 0)}{P(A_i - B_i \neq 0)} = \frac{a(1-b)}{a(1-b) + b(1-a)}$$

(b) Note when $a > b, p > 1/2$. Consider an up p down $1-p$ simple random walk and a gambler with M dollars and $N = 2M$. We have

$$\alpha_M = \frac{(q/p)^M - (q/p)^{2M}}{1 - (q/p)^{2M}} = \frac{1}{1 + (p/q)^M} = \frac{1}{1 + \left[\frac{a(1-b)}{b(1-a)}\right]^M}$$

Note when $a > b, a(1-b) > b(1-a)$ and thus $\alpha_M \to 0$ as $M \to \infty$.