MA3676 - Stochastic Models

1720996

March 26, 2021

Contents

1 Week 1

1.1 Axioms of Probability and conditional probabilities

1.1.1 Axioms of probability

- **Axiom** 1: To each event E there corresponds a number $\mathbb{P}[E] \geq 0$, where $\mathbb{P}[E]$ is called the probability of E.
- **Axiom** 2: $\mathbb{P}[\mathbb{S}] = 1$.
- **Axiom 3**: If events A_1, A_2, $A_3, \ldots$ are mutually exclusive then

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]. \tag{1}$$

1.1.2 Conditional probability

Definition 1

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}. \tag{2}$$

Definition 2 *Two events A and B are independent if* $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$.

Definition 3

$$\mathbb{P}[A_i|B] = \frac{\mathbb{P}[A_i]\mathbb{P}[B|A_i]}{\sum_{j+1}^{n} \mathbb{P}[A_j]\mathbb{P}[B|A_j]}. \tag{3}$$

2 Week 2

2.1 Discrete random variables

2.1.1 Bernoulli distribution

Definition 4 *If the probability of event E is p, then the variable I_E is an indicator variable, which is said to follow the Bernoulli distribution with parameter P*

$$\mathbb{P}[I_E = 1] = p, \, \mathbb{P}[I_E = 0] = 1 - p, \tag{4}$$

or in shorthand, $\mathbb{P}[1] = p$, $\mathbb{P}[0] = 1 - p$.

2.1.2 Independence of random variables

Definition 5 *Two random variables X and Y are independent if for any possible value x of X and any possible value y of Y the events $\{X = x\}$ and $\{Y = y\}$ are independent.*

2.1.3 Binomial distribution

Definition 6

$$\mathbb{P}[\{Y_n = m\}] = \binom{n}{m} p^m 1^{n-m}. \tag{5}$$

2.1.4 Poisson distribution

Definition 7

$$P(m) = \frac{\lambda^m}{m!} e^{-\lambda}. \tag{6}$$

2.1.5 Geometric Distribution

Definition 8

$$P(n) = p^{n-1}(1-p). \tag{7}$$

2.2 Sums and Series

2.2.1 Finite geometric sum

Definition 9

$$S_n = \sum_{k=0}^{n} a^k = 1 + a + a^2 + \cdots + a^n, \tag{8}$$

hence

$$S_n = \frac{1-a^{n+1}}{1-a}. \tag{9}$$

2.2.2 Convergence of series

In order for the series $\sum_{n=1}^{\infty} b_n$ to converge, the limit $\lim_{A\to\infty}(b_1 + b_2 + b_3 + \cdots + b_A)$ must exist.
A necessary condition for the convergence of this series is that the limit of the sequence b_n exists and is equal to 0:

$$\lim_{n\to\infty} b_n = 0 \tag{10}$$

2.2.3 Series examples

Definition 10 *Geometric series:*

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}. \tag{11}$$

Definition 11 *Power series:*

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x. \tag{12}$$

Definition 12 *Logarithmic series:*

$$\sum_{n=1}^{\infty} \frac{x^n}{n} = -ln(1-x). \tag{13}$$

Definition 13 *Trigonometric series:*

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = \sin x, \tag{14}$$

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = \cos x. \tag{15}$$

3 Week 3

3.1 Expectation value

3.1.1 Expectation of a random variable

Definition 14 *Consider a discrete random variable X which takes its values in a set $\{x_1, x_2, \ldots\}$ with probabilities $\mathbb{P}[\{X = x_i\}] = p_i$. Then, the expectation value of X is defined as*

$$\mathbb{E}[X] = \sum_i x_i p_i. \tag{16}$$

3.1.2 Expectation of functions of random variables

Definition 15

$$\mathbb{E}[g(X)] = \sum_i g(x_i) p_i. \tag{17}$$

3.1.3 Linearity of expectation

Definition 16

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]. \tag{18}$$

3.2 Expectation and variance values

3.2.1 Variance of a random variable

Definition 17

$$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]. \tag{19}$$

3.2.2 Variance of a binomial random variable

Definition 18

$$Var[\sum_{i+1}^{n} X_i] = n\, Var[X]. \tag{20}$$

4 Week 4

4.1 Generating functions

4.1.1 The concept of a stochastic process: terminology

- The collection of all indices marking the random variables: the index set.
- All possible values of each X_i: the state space.
- Particular values x_i taken by X_i are the states of the process.
- The elements of the index set are conventionally called times.
- The whole collection of values $\{x_i\}_{i=1}^{N}$ is a realisation of the process run for N steps.
- A stochastic process can be viewed as a single, multi-component random variable.

4.1.2 Definitions and basic properties

Definition 19 *A (discrete) "Random Walk" is a stochastic process S_n which can be represented as a sum of random variables ("steps"):*

$$S_n = S_0 + \sum_{k=1}^{n} X_k. \tag{21}$$

Definition 20 *A "Simple Random Walk" is obtained if X_k are independently identically distributed (i.i.d.) modified Bernoulli random variables taking values in the set $\{-1,\, 1\}$, i.e.*

$$\mathbb{P}[X_k = 1] = p, \quad \mathbb{P}[X_k = -1] = 1, \quad p + q = 1, \quad \forall k \in \mathbb{N}.$$

4.1.3 Generating functions: definition

Definition 21 *Consider a discrete random variable X, and denote $\mathbb{P}[X = n] = p_n$. We now define the corresponding generating function $G(s)$ as*

$$G(s) = \mathbb{E}[s^X]. \tag{22}$$

Definition 22

$$G(s) = \sum_{n} p_n s^n. \tag{23}$$

4.1.4 Expectation of a generating function

Definition 23

$$\mathbb{E}[X] = G'(1) \tag{24}$$

4.1.5 Variance of a generating function

Definition 24

$$Var[X] = G''(1) + G'(1) - G'(1)^2 \tag{25}$$

5 Week 5

5.1 Random Walks

5.1.1 Probability mass function

Definition 25

$$G_{S_n} = \sum_{m=0}^{n} \binom{n}{m} p^{n-m} q^m s^{n-2m}, \tag{26}$$

where the powers of s correspond to values of S_n.

Definition 26 *The probability of displacement by k is denoted as*

$$P_k^{(n)} \equiv \mathbb{P}[S_n - S_0 = k] = \binom{n}{\frac{n-k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}. \tag{27}$$

5.1.2 General random walks

Theorem 1 *Suppose $S_n = \sum_{k=1}^{n} X_k$ is a random walk such that X_k's are i.i.d. random variables, characterised by a generating function $G_X(s) = \mathbb{E}[s^X]$. Then, the generating function of S_n is*

$$G_{S_n}(s) = G_X(s)^n \tag{28}$$

5.1.3 Returning to the initial position

Theorem 2 *For an unrestricted unbiased one-dimensional random walk the probability of return $\mathcal{R} = 1$.*

Here, the number of returns N to the state i is represented by $\mathbb{E}[N]$.

Lemma 1

$$E[N] = \sum_{n=1}^{\infty} P_0^{(n)}. \tag{29}$$

Lemma 2 *If $\mathcal{R} < 1$, then*

$$\mathbb{E}[N] = \frac{\mathcal{R}}{1 - \mathcal{R}}. \tag{30}$$

6 Week 6

6.1 Gambler's ruin

Denote events such that:

- R_n: player A is eventually ruined starting from initial capital n.
- W_n: player A wins the first game starting from initial capital n.
- $L_n = \overline{W}_n$: player A loses the first game starting from initial capital n.

6.1.1 Probability of ruin

Definition 27 *The probability of ruin of a player from initial capital n is*

$$\mathbb{P}[R_n] = \mathbb{P}[R_n|W_n]\mathbb{P}[W_n] + \mathbb{P}[R_n|L_n]\mathbb{P}[L_n], \tag{31}$$

which simplifies to

$$P_n = pP_{n+1} + qP_{n-1}, \tag{32}$$

where $\mathbb{P}[R_n|L_n] = P_{n-1}$, $\mathbb{P}[R_n|W_n] = P_{n+1}$, $\mathbb{P}[W_n] = p$ and $\mathbb{P}[L_n] = q = 1 - p$.

Definition 28 *The probability of ruin from initial capital a with total pool $a + b$ is given by*

$$P_a = \frac{1 - (1/p)^b}{(p/q)^a - (1/p)^b}, \tag{33}$$

where $p \neq q$. If $p = q = 1/2$, then

$$P_a = \frac{b}{a + b}. \tag{34}$$

7 Week 7

7.1 First step decomposition

7.1.1 Time to absorption

Definition 29 *Let T_n be the random variable equal to the number of steps, starting from n, which the random walker will make before reaching wither of the boundaries. Denote as*

$$\mathcal{T} = \mathbb{E}[T_n]$$

the expected number of steps until the walker reaches either of the boundaries.

Definition 30

$$\mathcal{T}_n = A + B(q/p)^n + \frac{n}{q - p}, \qquad p \neq q, \tag{35}$$

$$\mathcal{T}_n = A + Bn - n^2, \qquad q = p = 1/2. \tag{36}$$

8 Week 8

8.1 Boundary Conditions

Use first step decomposition on the boundary conditions to solve for A and B. Use examples to understand.

9 Week 9

9.1 Introduction to Markov chains

Definition 31 *A stochastic process S_n with discrete state space and discrete index set is a Markov chain if $\mathbb{P}[S_{n+1} = j|S_0 = i_0, S_1 = i_1, \ldots, S_n = i_n] = \mathbb{P}[S_{n+1} = j|S_n = i_n]$, $\forall j, i_0, i_1, \ldots, i_n$ and $\forall n \in \mathbb{N}$.*

This says that the future (S_{n+1}) is conditionally independent of the past ($\{S_i\}_{i+1}^{n-1}$), given the knowledge of the present S_n.

9.2 Markov chains

9.2.1 Random walk on a circle

Definition 32 *Treat p_{ij} as elements of an $(n \times n)$-matrix and $\mathbf{p}_i(n)$ as elements of an n-component row vector $\underline{\pi}(n)$ where the matrix $\mathbf{p}$ provides the full description of transitions and $\underline{\pi}$ the current states of each position. The total probability decomposition is a product of these such that:*

$$\underline{\pi}(n + 1) = \underline{\pi}(n)\mathbf{p}. \tag{37}$$

9.2.2 Chapman-Kolmogorov equation

Theorem 3

$$\underline{\pi}(n) = \underline{\pi}(0)\mathbf{p}^n. \tag{38}$$

9.2.3 Multiplying stochastic matrices

Theorem 4 *A product of two (or more, by induction) stochastic matrices is again a stochastic matrix.*

10 Week 10

10.1 Markov chains

10.1.1 Equilibrium distribution

Theorem 5 *If such a distribution exists, then it means formally the existence of the limit*

$$\lim_{n\to\infty} \underline{\pi}(n) = \underline{\pi}, \tag{39}$$

and therefore

$$\underline{\pi} = \underline{\pi}\mathbf{p}. \tag{40}$$

10.1.2 Matrix analysis of Markov chains

Theorem 6 $\lambda = 1$ *is an eigenvalue of any stochastic matrix* $\mathbf{p}$.

11 Week 17

11.1 Markov chains

11.1.1 Revisiting a position

In a random walk with only ± 1 steps allowed, even positions can be reached with an even number of steps and vice versa. Thus a position i can be revisited only after an even number of steps:

$$P_{ii}^{(2m)} = \frac{(2m)!}{(m!)^2} p^m q^m \quad \text{and} \quad P_{ii}^{(2m-1)} = 0 \tag{41}$$

for any initial position.

Definition 33 *A state, return to which is only possible after* m, $2m$, $3m, \ldots$ *steps is called periodic if* $m \geq 2$. *Any other state is called aperiodic.*

Definition 34 *If the probability of return to a given state is equal to one, such a state is called recurrent. If this probability is less than one, the state is called transient.*

Note that this definition refers to the eventual probability of return, after an arbitrary number of steps.

11.1.2 Classification of states

These concepts apply to general discrete stochastic processes; in Markov chain applications $P_{ij}^{(m)}$ is $(p^{(m)})_{ij}$.

Definition 35 *A state i is said to communicate with a state j if it is possible eventually to reach j starting from i, i.i., if $P_{ij}^{(m)} > 0$ for some m.*

Definition 36 *States i and j intercommunicate if, in addition,* $\exists m' : P_{ji}^{(m')} > 0$.

This property is denoted as $i \leftrightarrow j$, and it is transitive:

$$\text{If } i \leftrightarrow j \text{ and } j \leftrightarrow k, \text{ then } i \leftrightarrow k. \tag{42}$$

11.1.3 Recurrence time

Starting from a state i the probability that the chain will next visit the same state after m steps can be formally expressed ass

$$f_{ii}^{(m)} = \mathbb{P}[(X_{n+m} = i \cap (X_{n+k} \neq i \forall k : 1 \leq k \leq m-1)|X_n = i)]. \tag{43}$$

The sum

$$R_i = \sum_{m=1}^{\infty} f_{ii}^{(m)} \tag{44}$$

is the probability of return, so $R_i = 1$ for a recurrent state i. This means that $f_{ii}^{(m)}$ has the meaning of a discrete probability distribution of recurrence "times" $T_i : f_{ii}^{(m)} = \mathbb{P}[T_i = m]$.
Thus $\mu_i = \mathbb{E}[T_i] = \sum_{m=1}^{\infty} m f_{ii}^{(m)}$ is the expectation of the recurrence time - often known as the mean recurrence time for a state i.

Definition 37 *Depending on whether μ_i is finite or infinite, all recurrent states are classified as positive-recurrent, or null-recurrent, respectively.*

Definition 38 *A positive-recurrent aperiodic state is ergodic.*

Two more special types of states are:

- **Ephemeral** states: those states i for which $p_{ki} = 1$, for any k, and
- **Absorbing** states: those states i for which $p_{ii} = 1$.

11.1.4 States are "contagious"

Theorem 7 *Two intercommunicating states must be of the same type.*

Theorem 8 *Let i be a recurrent state which communicates with j; then i and j intercommunicate, and j is also a recurrent state of the same type and period as i.*

Corollary 1 *A pair of recurrent states of the same type either intercommunicate, or neither of them communicates with the other.*

Corollary 2 *A recurrent state of one type cannot communicate with a recurrent state of another type, nor with a transient state.*

Note, a transient state can communicate with other transient states, as well as with both types of recurrent states.

11.1.5 Closed sets of states

Definition 39 *A set C of states is closed if no state in C communicates with any state outside of C. A sufficient condition for this is*

$$p_{ij} = 0 \forall i \in C \text{ and } \forall j \notin C. \tag{45}$$

Definition 40 *A closed set C of states is irreducible if every pair of states in C intercommunicate.*

- Since a recurrent state cannot communicate with a transient state, the set of all recurrent states is closed, but not necessarily irreducible.
- If a closed set contains only one state then that state is absorbing.
- An irreducible closed set contains no smaller closed set.

Theorem 9 *The state space $\mathbb{S}$ of a Markov chain can be uniquely partitioned as*

$$\mathbb{S} = T \cup C_1 \cup C_2 \cup \ldots$$

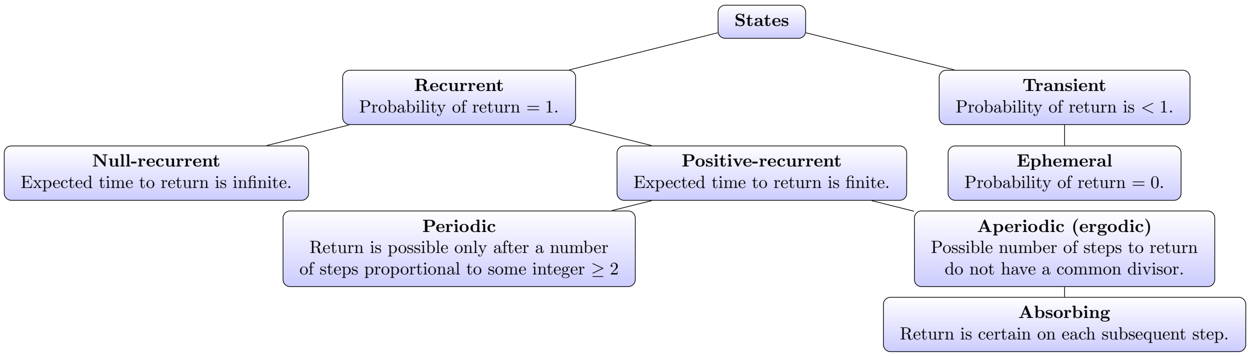
where T is the set of all transient states, and C_1, $C_2, \ldots$ are irreducible closed sets of recurrent states, each set containing states of the same type.

Theorem 10 *If the state space is finite, then at least one state is recurrent, and, assuming the chain is time-homogeneous, all recurrent states are positive recurrent.*

11.1.6 Corollaries of the decomposition theorem

Corollary 3 *If the chain starts in, or eventually reaches one of the irreducible closed sets C_r, then it never leaves it. C_r thus becomes the state space for the subsequent part of the realisation of the chain, and every state in C_r is visited infinitely often. Thus the subset C_r, together with the corresponding transition probabilities, forms a Markov chain in it's own right.*

Corollary 4 *If the chain starts in the set T of transient states, then it either moves between states within T forever, visiting any one state only a finite number of times, or it eventually enters one of the closed sets C_r of recurrent states, and it remains in C_r forever after.*



11.1.7 Block matrices

$$M = \left(\begin{array}{ccc|cc|cc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ \hline 15 & 16 & 17 & 18 & 19 & 20 & 21 \\ 22 & 23 & 24 & 25 & 26 & 27 & 28 \\ 29 & 30 & 31 & 32 & 33 & 34 & 35 \end{array}\right) = \left(\begin{array}{ccc} A & B & C \\ D & E & F \end{array}\right).$$

Each of the blocks is defined in the obvious way:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 8 & 9 & 10 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 5 \\ 11 & 12 \end{pmatrix} \quad \text{etc.}$$

If we were to then have a matrix N such that

$$N = \begin{pmatrix} G & H \\ J & K \\ L & S \end{pmatrix}$$

that is the correct size so MN exists, then

$$MN = \begin{pmatrix} A & B & C \\ D & E & F \end{pmatrix} \begin{pmatrix} G & H \\ J & K \\ L & S \end{pmatrix} = \begin{pmatrix} AG + BJ + CL & AH + BK + CS \\ DG + EJ + FL & DH + EK + FS \end{pmatrix}. \tag{46}$$

11.1.8 Permutations of rows and columns

Consider a 3×3-transition matrix, corresponding to states 1, 2 and 3:

$$\mathbf{p} = \begin{pmatrix} .1 & .7 & .2 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{pmatrix}.$$

If we now decide to list the states in the order 3, 2, 1, then we need to swap the first and third row, **and** the first and third column, so that we get

$$\begin{pmatrix} .4 & .4 & .2 \\ .2 & .5 & .3 \\ .2 & .7 & .1 \end{pmatrix}.$$

11.1.9 Reducibility

Theorem 11 *If columns and rows of a square matrix A can be permuted so that it acquires the form*

$$A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \tag{47}$$

where B (and hence D) are square matrices, and 0 is a matrix consisting of all zeros, then A is called reducible.

Theorem 12 *Transition matrix* $\mathbf{p}$ *of a decomposable (reducible) Markov chain is reducible.*

Suppose $\mathbb{S} = C \cup T$, with k_C states in the closed set, and k_T transient states. By definition of reducibility, the transition matrix $\mathbf{p}$ can be rearranged into the form

$$\mathbf{p} = \begin{pmatrix} P & 0 \\ R & Q \end{pmatrix}. \tag{48}$$

Here, P corresponds to transitions among recurrent states of the closed set C, elements of Q are probabilities of transitions among transient states, and elements of R give the probabilities of transitions from transient to recurrent states. This form is retained after an arbitrary number of steps.

12 Week 18

12.1 Markov Chains

12.1.1 Perron-Frobenius theorem

Definition 41 *If all other eigenvalues λ satisfy the strict inequality $|\lambda| < r$, then A is called primitive.*

Theorem 13 *Suppose a matrix A is irreducible, and all elements of A are non-negative. Then the following statements are true:*

1. *A has at least one real positive eigenvalue.*
2. *The largest (if there is more than one) such real positive eigenvalue r dominates all other eigenvalues: If α is any other eigenvalue of A, then $|\alpha| \leq r$.*
3. *r is non-degenerate, i.i., it is a simple root of the characteristic equations*

$$\det(A_\lambda \mathbb{I}) = 0. \tag{49}$$

4. *Corresponding to r are the right and left eigenvectors all of whose components are positive. Conversely, the only eigenvectors whose components are positive are those corresponding to r.*
5. *r increases when any element of A increases. Correspondingly, r decreases when any element of A decreases.*
6. *If A is not primitive, i.e., there are $d > 1$ eigenvalues, such that $|\lambda| = r$, then they are all simple and different, and are located symmetrically on the unit circle in the complex plane: if they are numbered from 0 to $d-1$, they can be written as*

$$\lambda_k = re^{2\pi ik/d}, \tag{49}$$

with $k = 0, 1, \ldots, d-1$. If there are two such roots, they are equal to r and $-r$.
Further, if $d > 1$, the rows and columns of the matrix can be permuted to bring it to the block form

$$\begin{pmatrix} 0 & A_1 & 0 & \cdots & 0 \\ 0 & 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_{d-1} \\ A_d & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

12.1.2 Limiting behaviour

Corollary 5 *If A is irreducible and primitive, and x and y are the right (column) and left (row) eigenvectors of A respectively, corresponding to the dominant eigenvalue r, and normalised so that $yx = 1$, then*

$$\lim_{n\to\infty} (1/r)^n A^n = x \otimes y, \tag{49}$$

where $\otimes$ is used to denote the "outer" product of vectors: column times row.

Such products are known as dyadic products and the limits are uniform for all elements, so

$$\lim_{n\to\infty} (1/r)^n (A^n)_{ij} = x_i y_i.$$

Theorem 14 *For irreducible and primitive* $\mathbf{p}$,

$$\lim_{n\to\infty} \mathbf{p}^n = e \otimes \pi. \tag{50}$$

It follows that

Corollary 6 *If* $\mathbf{p}$ *is irreducible and primitive, there exists an equilibrium distribution.*

12.1.3 Reducible matrices

Given a reducible matrix A, its rows and columns can be permuted to the canonical form:

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix}.$$

If B is reducible in its turn, its rows and columns can be permuted further to exhibit the structure:

$$\begin{pmatrix} B_{11} & 0 & 0 \\ B_{21} & B_{22} & 0 \\ C_1 & C_2 & D \end{pmatrix}, \tag{51}$$

where $\begin{pmatrix} C_1 & C_2 \end{pmatrix} = C$.
For the application to Markov chains, only the following special case of the canonical form is relevant:

$$p = \begin{pmatrix} P_1 & 0 & \cdots & 0 & 0 \\ 0 & P_2 & \cdots & 0 & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & P_n & 0 \\ & & R & & Q \end{pmatrix}. \tag{52}$$

Each block P_k contains probabilities of transitions inside k-th closed set, the Q block describes transitions among transient states, and the R block describes transitions from transient to recurrent states.

Theorem 15 *If the chain has transient states and/or more than one closed set of recurrent states, then p is reducible, and vice versa.The number of closed sets of recurrent states is equal to the multiplicity of the dominant eigenvalue $r = 1$.*

12.1.4 Classification of chains

Considering only chains with a finite overall number of states:

- **Reducible:**
  - One closed set of recurrent states:
    Single eigenvalue equal to 1.
    Equilibrium distribution exists.
  - Two or more closed sets of recurrent states:
    Two or more eigenvalues equal to 1.
    Equilibrium distribution does not exist.

- **Irreducible:**
  - **Periodic:**
    Single eigenvalue equal to 1. Two or more eigenvalues with absolute value equal to 1. Period equal to the number of such eigenvalues.
    Equilibrium distribution does not exist.
  - **Ergodic:**
    Single eigenvalue equal to 1. No other eigenvalues with absolute value equal to 1.
    Equilibrium distribution exists.

12.1.5 Algebraic test of reducibility

Definition 42 *A matrix M is said to be greater than 0 ($M > 0$) if each element of M is greater than 0 ($M_{ij} > 0 \forall i, j$).*

Theorem 16 *An $n \times n$-matrix p is reducible iff $(\mathbb{I}_n + p)^{n-1} > 0$.*

12.1.6 Long-time limit

$$p^n = \begin{pmatrix} P^n & 0 \\ V_n & Q^n \end{pmatrix}, \tag{53}$$

where $V_n = RP^{n-1} + QRP^{n-2} + Q^2RP^{n-3} + \ldots + Q^{n-2}RP + Q^{n-1}R$. This sum can be written in compact notation as:

$$V_n = \sum_{j=0}^{n-1} Q^j RP^{n-1-j}. \tag{54}$$

V_n satisfies a recursion:

$$V_n = QV_{n-1} + RP^{n-1}. \tag{55}$$

12.1.7 Evaluating the limits

- $\lim_{n\to\infty} P^n = \begin{pmatrix} \Pi & 0 \\ V & 0 \end{pmatrix}$, where $V = \lim_{n\to\infty} V_n$
- $\lim_{n\to\infty} Q_n = 0$.

Theorem 17

$$R\Pi + QV = V \tag{56}$$

$$V = (\mathbb{I} - Q)^{-1} R\Pi, \tag{57}$$

where $\mathbb{I}$ is the unit matrix of the same dimension as Q.

12.1.8 Special case

Theorem 18 *When the chain has a single closed set of recurrent states:*

$$\lim_{n\to\infty} p^n = e_{r+t} \otimes (\pi, 0_t). \tag{58}$$

Here, t is the number of transient states and r is the number of columns in R.

12.1.9 Absorption problems in reducible chains

Theorem 19 *A Markov chain with a single ergodic closed set possesses an equilibrium distribution.*

13 Week 19

13.1 Markov chains

13.1.1 Absorption probabilities

Let the transition matrix p be of the form

$$p = \begin{pmatrix} P & 0 \\ R & Q \end{pmatrix},$$

where P has a block-diagonal structure with each square block containing elements encoding transitions among the states of the n-th closed set.
The structure of R is

$$R = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1r} \\ R_{21} & R_{22} & \cdots & R_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ R_{t1} & R_{t2} & \cdots & R_{tr} \end{pmatrix}.$$

Each $R_{\alpha n}$ is a row of k_n elements $R_{\alpha j}$ describing the transitions between the α-th transient state and the j-th state in the closed set C_n.
If we are interested onlny in the properties of the chain up to the time of absorption, the subsequent transitions among the states of a closed set are irrelevant. Thus, we can replace our original chain with an equivalent one which:

$$\text{Each of the closed sets is replaced by a single compound absorbing state.} \tag{59}$$

Correspondingly, each row $R_{\alpha j}$ is replaced with the number

$$\tilde{R}_{\alpha n} = \sum_{j \in C_n} R_{\alpha j}. \tag{60}$$

The probabilities of absorption are given by the elements of the matrix

$$\tilde{V} = (\mathbb{I} - Q)^{-1} \tilde{R}. \tag{61}$$

13.1.2 First-step decomposition in absorption problems

Let $V_{\alpha j}$ be the eventual probability to be in the j-th recurrent state starting from the transient state α. Then:

Theorem 20

$$V_{\alpha j} = \sum_{r \in C_j} \Pi_{rj} R_{\alpha r} + \sum_{\beta} V_{\beta j} Q_{\alpha \beta}. \tag{62}$$

In matrix form, this is the same as eq.(57).

13.1.3 Recurrence time

Theorem 21 *The mean recurrence time for an ergodic state j is equal to the inverse of the equilibrium probability in the state j.*

Theorem 22 *The mean recurrence time μ for a state j is given by:*

$$\mu_j = \sum_{m=1}^{\infty} m f_{jj}^{(m)} = \frac{1}{\pi_j}. \tag{63}$$

14 Week 20

14.1 Markov chains

14.1.1 Mean time to absorption

Theorem 23

$$\mu = (\mathbb{I} - Q)^(- 1) e^{(t)}, \tag{64}$$

where μ is the column of mean absorption times and $(\mathbb{I} - Q)^(- 1)$ is the mean number of visits.

15 Week 21

15.1 Branching Processes

15.1.1 Population dynamics

- **Deterministic**: In deterministic models the rules of the evolution of a population size from one generation to the next are fixed, and complexity arises, e.g., from non-linear interaction between different subspecies of the population (e.g., predators eat prey).
- **Stochastic**: In stochastic models one takes into account the fact that each member of the population may have a varying number of immediate descendants, so that this number is modeled as a random variable subject to a certain probability law.

A deterministic model often arises as the average behaviour of a stochastic model when the number of individuals is large.

15.1.2 Galton-Watson branching process

Theorem 24 *The total population size of the $n+1$-st generation is*

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i. \tag{65}$$

The X_i are taken to be i.i.d. random variables.
Define two types of generating functions. $\mathbb{G}_n(s) = \mathbb{E}[s^{Z_n}]$ for the population in the n-th generation. and $G(s) = \mathbb{E}[s^X]$ for family size.

Theorem 25

$$\mathbb{G}_{n+1}(s) = \mathbb{G}_n[G(s)] \tag{66}$$

and

$$\mathbb{G}_n(s) = G(\mathbb{G}_{n-1}(s)). \tag{67}$$

Theorem 26 *The probability of eventual extinction ξ is given by*

$$\xi = \lim_{n\to\infty} \mathbb{G}_n(0). \tag{68}$$

This also means

$$\xi = G(\xi). \tag{69}$$

Corollary 7 *The extinction probability ξ of the branching process is given by the smallest of the two non-negative roots of the equation*

$$\xi = G(\xi),$$

where $G(s)$ is the generating function for the distribution of the number of descendants of a single individual.

Corollary 8 *Extinction is certain iff mean "family" size is less than or equal to one.*

16 Week 23

16.1 Introduction to Martingales

16.1.1 Martingales

Definition 43 *A martingale is a stochastic process which possesses the property of being a "fair game", and does not allow for infinite expectations.*
Formally: Z_n is a martingale with respect to itself if:

1. $\mathbb{E}[Z_{n+1}|\{Z_i\}_{i=1}^n] = Z_n$ *and*
2. $\mathbb{E}[|Z_n|] < \infty$.

A simple unbiased random walk without boundaries is a martingale.

16.1.2 Biased random walk

1. A more general definition of martingale with respect to another process (as opposed to with respect to itself) could be Y_n is a martingale with respect to S_n.
2. Since martingales have many very useful properties, constructing a martingale on the basis of the process one wants to study is a powerful mathematical tool.

16.1.3 Properties of martingales

Theorem 27 *If Z_n is a martingale, then*

$$\mathbb{E}[Z_n] = \mathbb{E}[Z_0]. \tag{70}$$

16.1.4 De Moivre's martingale

Definition 44 *De Moivre's martingale is defined as*

$$Y_n = (q/p)^{S_n}, \tag{71}$$

where $p \neq q$.

16.1.5 Stopping times

Definition 45 *A random variable T is a stopping time with respect to a stochastic process Y_n if $\{T = n\} \in \mathcal{F}$ where $\mathcal{F}_n$ is the powerset associated with the sample space of all possible realisations of the sequence*

$$Y_0,\, Y_1,\, Y_2, \ldots,\, Y_{n-1},\, Y_n.$$

$\mathcal{F}_n$ is known as a filtration. Martingales are properly defined with respect to filtrations, rather than the underlying stochastic processes:

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = Y_n \tag{72}$$

defines a martingale.

17 Week 24

17.1 Introduction to martingales

17.1.1 Optional stopping theorem

Theorem 28 *Let Y_n be a martingale with respect to a filtration $\mathcal{F}_n$ and let T be a stopping time. Then, $\mathbb{E}[Y_T] = \mathbb{E}[Y_0]$ if the following conditions are satisfied:*

1. $\mathbb{P}[T < \infty] = 1$,
2. $\mathbb{E}[T] < \infty$,
3. *$\exists$ a constant $\rfloor$ such that for all $n < T$ the following condition is satisfied:*

$$\mathbb{E}\left[|Y_{n+1} - Y_n|\middle|\mathcal{F}_n\right] \leq c. \tag{73}$$

17.1.2 Gambling strategies

Consider a martingale $W_n = \sum_{k=1}^{n} Z_k$. Suppose now the gambler attempts to execute a strategy by varying the amount of individual bets depending on the outcome(s) preceding game(s). Let us denote these bets as α_n, so that the capital of the gambler after n games is

$$C_n = \sum_{k=1}^{n} \alpha_k Z_k. \tag{74}$$

Theorem 29 *If each bet α_k depends only on the outcomes of the preceding $k-1$ games, then C_n is a martingale.*

17.1.3 Wald's identity

Assuming the validity of the optional stopping theorem, we have for a suitable stopping time T,

$$\mathbb{E}[s^{S_T} G(S)^{-T}] = 1. \tag{75}$$