

# Contents

<b>1</b>	<b>Probability</b>	<b>2</b>
1.0.0.1	Sample Space and Outcome . . . . .	2
1.0.0.2	Event . . . . .	2
1.0.0.3	Disjunction of Events . . . . .	2
1.0.0.4	Conjunction of Events . . . . .	3
1.0.0.5	Mutually Exclusive Events . . . . .	3
1.0.0.6	Axioms of Probability . . . . .	3
1.0.0.7	Properties . . . . .	4
1.0.0.8	Conditional Probability . . . . .	4
1.0.0.9	Joint Probability . . . . .	4
1.0.0.10	Marginal Probability . . . . .	5
<b>2</b>	<b>Stochastic Processes</b>	<b>5</b>
2.0.1	Classifications . . . . .	5
2.0.1.1	State Space . . . . .	5
2.0.1.2	Discrete State Space . . . . .	5
2.0.1.3	Continuous State Space . . . . .	5
2.0.1.4	Time Parameter . . . . .	5
2.0.1.5	Discrete Time Parameter . . . . .	6
2.0.1.6	Continuous Time Parameter . . . . .	6
2.1	Discrete State Space and Time . . . . .	6
2.1.0.7	Markov Chain . . . . .	6
2.1.0.8	Homogeneous Markov Chains . . . . .	6
2.1.0.9	State Transistion Diagram . . . . .	7
2.1.0.10	Chapman-Kolmogorov Equation . . . . .	7
2.1.0.11	Irreducible Markov Chain . . . . .	7
2.1.0.12	Recurrent State . . . . .	8
2.1.0.13	Recurrent Non-null . . . . .	8
2.1.0.14	Recurrent Null . . . . .	8
2.1.0.15	Periodicity . . . . .	8
2.1.0.16	State Probability . . . . .	9
2.1.0.17	Steady State Therom . . . . .	9
2.1.0.18	Residence Time in a State . . . . .	10
2.2	Continuous Time . . . . .	10
2.2.0.19	Assumption: Continuos Time Markov Chain . . . . .	11
2.2.0.20	Transistion Probability . . . . .	11

2.2.0.21	Apply Chapman-Kolmogorov Equation . . . .	11
2.2.0.22	Partial of $P_{ij}(v, t)$ . . . . .	12
2.2.0.23	Rate Definitions for Simplificaiton . . . . .	12
2.2.0.24	State Probabilitiy . . . . .	12
2.2.0.25	Steady State Results . . . . .	13
2.2.0.26	Residence Time in State . . . . .	13
<b>3</b>	<b>Derivations</b>	<b>13</b>
3.1	Probability . . . . .	13
3.2	Stochastic Processes . . . . .	13
3.2.0.27	Chapman-Kolmogorov Equation . . . . .	13
3.2.0.28	Partial of $P_{ij}(v, t)$ . . . . .	13
3.2.0.29	Derivative of $\pi_j(t)$ . . . . .	13

The book is optimized for looking up facts. However, it contains pointers to the end of the books that give proof. I am debating whether or not to move the examples into a separate section as well. For now they are kept with the facts.

## 1 Probability

**1.0.0.1 Sample Space and Outcome** We perform random experiments and the sample space is the set of possible outcomes. For example, consider rolling a die. The set of possible outcomes are:

$$S = \{1, 2, 3, 4, 5, 6\}$$

**1.0.0.2 Event** An event is a subset of the sample space. An example event is rolling a die and getting an even odd outcome:

$$E = \{1, 3, 5\}$$

**1.0.0.3 Disjunction of Events** The event  $E$  occurs if  $E_1$  or  $E_2$  occur. Another way to imagine this is the union of events:  $E = E_1 \cup E_2$ .

**1.0.0.4 Conjunction of Events** The event  $E$  occurs if  $E_1$  and  $E_2$  occur. Another way to imagine this is the intersection of events:  $E = E_1 \cap E_2$ . Some alternative ways of writing this are:

$$P(E) = P(E_1 \cap E_2)$$

$$P(E) = P(E_1 \wedge E_2)$$

$$P(E) = P(E_1, E_2)$$

$$P(E) = P(E_1 E_2)$$

**1.0.0.5 Mutually Exclusive Events** Events  $E_1$  and  $E_2$  are mutually exclusive if only one of them can occur in a single experiment. For example, the event rolling an even number and the event rolling an odd number on a die are mutually exclusive events:

$$E_{\text{even}} \cap E_{\text{odd}} = \{2, 4, 6\} \cap \{1, 3, 5\} = \emptyset$$

**1.0.0.6 Axioms of Probability** These are the rules we accept as truth without proof. We build probability upon these axioms.

1.  $0 \leq P(E) \leq 1$ , for any event  $E$ . In the smallest case, the event cannot occur which is indicated by a probability of 0. In the largest case, the event always occurs, which is indicated by the probability of 1.
2.  $P(S) = 1$ , where  $S$  is the sample space. The sample space contains all possible outcomes for each experiment. It's reasonable to accept that an event from the sample space always occurs.
3. For a potentially infinite set of mutually exclusive events  $E_1, E_2, \dots$

$$P(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

It makes sense that events that do not share outcomes for a single event, can have their probabilities added to arrive at the probability of combining the outcomes from the events.

**1.0.0.7 Properties** From the above axioms, we get the following useful properties (TODO proof):

1. For any event  $E$ , let  $\overline{E}$  be the complement of  $E$ . More concretely,  $\overline{E} = S - E$ , where  $S$  is the sample space. Then  $E$  and  $\overline{E}$  are mutually exclusive.
2.  $P(\emptyset) = 0$  You can never get none of the outcomes of the sample space.
3. If  $E_1$  and  $E_2$  are mutually exclusive events then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1, E_2)$$

**1.0.0.8 Conditional Probability** We use conditional probability to model the probability given knowing some circumstance has happened. Given two event  $E$  and  $F$ , the conditional probability, the probability of  $F$  given  $E$  has occurred, is defined as ( $P(E) \neq 0$ ):

$$P(F|E) = \frac{P(E, F)}{P(E)}$$

An example is what is the probability of rolling a 3, given that we rolled an odd number. Let  $F = \{3\}$  and  $E = \{1, 3, 5\}$ :

$$P(F|E) = \frac{P(E, F)}{P(E)} = \frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$$

**1.0.0.9 Joint Probability** In queuing theory, we often have to use multiple sample sapce. The theory in this book so far has covered only a single probability space.

Suppose we have two sample spaces  $S_1$  and  $S_2$ . The outcomes of the joint probability space are the tuples that result from the cross product of the two sample spaces:

$$S_{joint} = S_1 \times S_2$$

For example, consider rolling a die and a coin

$$\begin{aligned}S_{die} &= \{1, 2, 3, 4, 5, 6\} \\S_{coin} &= \{H, T\} \\S &= S_{die} \times S_{coin} \\S &= \{(1, H), (2, H), (3, H), (4, H), (5, H), (6, H), \\&\quad (1, T), (2, T), (3, T), (4, T), (5, T), (6, T)\}\end{aligned}$$

**1.0.0.10 Marginal Probability** Given the joint probabilities, we might want to compute the probabilities of only one of the sample spaces.

For example, suppose that we know the joint probability of the number of jobs at server 1 and server two and we want to compute the probability of the number of jobs at server 1 only. We can apply Marginal probability to determine that.

## 2 Stochastic Processes

A family of random variables, indexed by time

### 2.0.1 Classifications

**2.0.1.1 State Space** The set of possible values (states).

**2.0.1.2 Discrete State Space** Example: the number of jobs in the system. ( $S = 0, 1, 2, 3, \dots$ ). We will only deal with the discrete case in this class. To make notation easier the state is usually identified by the number.

**2.0.1.3 Continuous State Space** Example: motion of a particle. We will not be studying this in this class.

**2.0.1.4 Time Parameter** There are two ways to observe times in Stochastic processes.

**2.0.1.5 Discrete Time Parameter** We consider the states at  $X_0, X_1, \dots, X_n$ , and so on. For example, looking at the state of the system at the  $i^{th}$  hour.

**2.0.1.6 Continuous Time Parameter** The states are function of time  $t, X(t)$ .

## 2.1 Discrete State Space and Time

There might be a dependency between the previous time interval  $X_i$ 's and the states those time interval can be in that need to be model in the current  $X_n$ .

$$P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0)$$

The number of dependency combinations is exponential because  $X_{n+1}$  depends on  $X_n$  to  $X_0$  and  $X_n$  depends on  $X_{n-1}$  to  $X_0$  and so on.

**2.1.0.7 Markov Chain** As a result of the exponential size, we make a simplifying assumption. We only use the latest information.  $X_{n+1}$  only depends on  $X_n$ . Now we are left with transition probabilities:

$$P(X_{n+1} | X_n) = P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0)$$

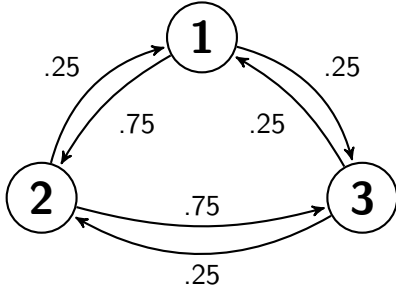
Given  $P(X_0 = i)$  for all  $i$ 's, we can compute any state. However, notice that the formula depends on  $n$ , the discrete time that has passed so far, which make analysis difficult still. For example, in the 9am one hour interval the number of jobs in a login system tends to be higher than at 2am.

**2.1.0.8 Homogeneous Markov Chains** We now make the assumption that the transition probabilities do not depend on time. For example, the transition probabilities for the number of webcrawling robots requesting a webpage remain the same despite the time. So we can write

$$P(X_{n+1} = j | X_n = i) = P(X_n = j | X_{n-1} = i), \forall n, i, j \geq 0$$

Which is abbreviated to  $P_{ij}$ , the probability of going from state  $i$  to state  $j$ .

**2.1.0.9 State Transition Diagram** It's a diagram that enumerates all possible state transitions, and annotates the edges with the probabilities of going from one state to the next. Below is an example of a state transition diagram.



**2.1.0.10 Chapman-Kolmogorov Equation** Let  $P_{ij}^{(m)}$  be the  $m$  step transition probability from state  $i$  to  $j$  defined as:

$$P_{ij}^{(m)} = P(X_{n+m} = j | X_n = i)$$

To take  $m$  steps to go from (possibly visiting  $m$  in an intermediate state multiple times), you can take  $m-1$  to steps. At step  $m-1$ , you can arrive at any state  $k$ . To go from each state  $k$  at the  $m-1$  step to the  $m$  step, you just apply the transition probability.

We can sum the probabilities because the probabilities of going from state  $k$  in the  $m-1$  step to state  $j$  are mutually exclusive (they are mutually exclusive because they are different states).

Lastly, we can multiply the probability of going to state  $k$  in  $m-1$  steps by the probability of going to state  $j$  because the probabilities are independent.

$$P_{ij}^{(m)} = \sum_k p_{ik}^{(m-1)} p_{kj}$$

Which is exactly the same as taking 1 step, and then  $m-1$  steps.

$$P_{ij}^{(m)} = \sum_k p_{ik} p_{kj}^{(m-1)}$$

For a derivation of this, see (3.2.0.27).

**2.1.0.11 Irreducible Markov Chain** allows every state to be reached from every other state for all pairs of states  $i$  and  $j$ . More concretely:

$$\forall i \forall j \neq i : \exists m_{ij} : P_{ij}^{(m_{ij})} > 0$$

**2.1.0.12 Recurrent State** : State  $j$  is recurrent if after leaving state  $j$  then you are guaranteed to eventually come back.

Let  $f_j^{(n)}$  be the probability that you first return to state  $j$  in  $n$  steps.

Notice that  $f_j^{(0)} = 0$  because it's impossible to come back without taking any steps. Also notice that  $f_j^{(1)} > 0$  is only possible if the transition pointing to itself  $P_{jj} > 0$ .

$j$  is recurrent if and only if

$$f_j = \sum_{n=1}^{\infty} f_j^{(n)} = 1$$

**2.1.0.13 Recurrent Non-null** A state  $j$  is recurrent non-null if and only if we get back to state  $j$ , but it does not take forever. More concretely:

$j$  is recurrent non-null if and only if  $j$  is recurrent ( $f_j = 1$ ) and

$$M_j = \sum_{n=1}^{\infty} n f_j^{(n)}$$

Where  $M_j$  is the expected number of steps to come back.

**2.1.0.14 Recurrent Null**  $j$  is recurrent null if and only if  $j$  is recurrent and  $j$  is not recurrent non-null.

TODO EXAMPLE WITH DIAGRAM

TODO EXAMPLE WITH DIAGRAM

**2.1.0.15 Periodicity** A state  $j$  is periodic if and only if the only way to come back to state  $j$  is to take  $r, 2r, 3r, \dots, cr$ , steps.

If a state  $j$  is not periodic, it's called aperiodic

If state  $j$  has a self loop ( $p_{jj} > 0$ ), then state  $j$  is aperiodic

If the system is an irreducible Markov Chain, and contains a self loop, then all states  $j$  are aperiodic.



**2.1.0.16 State Probability** Let  $X_n$  be the random variable for the state at interval  $n$ , then in a homogeneous Markov Chain we have that:

$$\pi_j^{(n)} = P(X_n = j) \text{ - at step } j$$

Let  $p_{ij}$  be the transition probability for going from state  $i$  to state  $j$  (independent of time, so it's the same for all intervals  $(X_n)$ , then

$$\pi_j^{(n+1)} = \sum_{i=0}^{\infty} p_{ij} \pi_i^{(n)} \text{ (by applying total probability)}$$

Given initial conditions  $\pi_j^{(0)} \forall j$ , we can compute all  $\pi_j^{(i)}$  apply the above formula recursively.

**2.1.0.17 Steady State Theorem** (Equilibrium Probability Theorem) (I made up this name. It seems better than 'Fundamental Theorem.)

If a homogeneous Markov Chain is irreducible and aperiodic, then there exists a limiting probability (equilibrium):

$$\pi_j = \lim_{n \rightarrow \infty} \pi_j^{(n)}$$

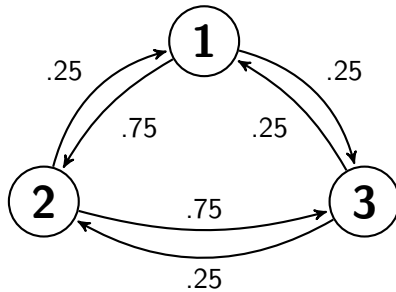
which is independent of the initial conditions  $\pi_j^{(0)}$ .

Moreover, if all states  $j$  are recurrent non-null, then  $\pi_j$  is non zero and can be uniquely determined from the equations:

$$\pi_j = \sum_i p_{ij} \pi_i, \forall j$$

$$\sum_j \pi_j = 1$$

#### 2.1.0.17.1 Example



Let the initial state be (1). This is the 'initial condition'

n	0	1	2	3	4	...	$\infty$
$\pi_1^{(n)}$	1	0	.25	.187	.203		.2
$\pi_2^{(n)}$	0	.25	.062	.359	.254		.28
$\pi_3^{(n)}$	0	.75	.688	.454	.543		.52

To get the steady state column, you have to solve the system of equations that results from the above theorem. From the table we can see that the empirical basis for the steady state theorem. As you compute more columns, it approaches the values computed from the steady state equations.

**2.1.0.18 Residence Time in a State** The residence time, is the amount of time that the system spends in state j for m steps, given that we are already in state j.

$$P(\text{in state j for m steps} | \text{already in state j}) \\ = P(\text{in state j for m steps} \wedge \text{step m+1 is not j} | \text{already in state j})$$

We have that the transition probabilities are independent of time (homogeneous), so we can multiple the individual probabilities.  $p_{jj}$  is the homogeneous transition probability of going from state j to state j.

$$= p_{jj}^m (1 - p_{jj})$$

## 2.2 Continuous Time

In continuous time, state transitions happen independent of clock ticks. We can extend the discrete time theory we have developed by considering the times  $t_0, t_1, t_2, \dots, t_n$ . These are the times when state transitions occur.

Note that, if we can accept accuracy within a time interval, then we can just use discrete time. Arrival and departure models generally use continuous time. This distribution is heavily connected with the exponential function.

**2.2.0.19 Assumption: Continuous Time Markov Chain** Let  $X(t)$  be a random variable, where the distribution of the random variable depends on the real parameter  $t$ . For instance, at  $t = 1.0$  the distribution of  $X$  might be exponential, but at  $t = \pi$  the distribution might be normal. Given the above notation for state transition, we have the following expression for the probability of the  $(n + 1)^{th}$  state transition.

$$P(X(t_{n+1}) = j | X(t_n) = i_n, \dots, X(t_0) = i_0)$$

Applying the Markov assumption that the next state transition only depends on the previous state gives us:

$$P(X(t_{n+1}) = j | X(t_n) = i_n)$$

All the definitions we created in discrete time for recurrent, irreducible, etc apply to continuous time as well by considering only the times when state transitions occur.

**2.2.0.20 Transition Probability** Let the transition probability be the probability of going from one state  $i$  at time  $v$ , to another state at time  $t$ .

$$P_{ij}(v, t) = P(X(t) = j | X(v) = i)$$

There is a special for the above formula:

$$P_{ij}(t, t) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

**2.2.0.21 Apply Chapman-Kolmogorov Equation** We can apply the Chapman-Kolmogorov Equation to the transition probability. The probability of going

1. from state  $i$  at time  $v$
2. to state  $j$  at time  $v$

is the same as going

1. from state  $i$  at time  $v$

2. to state  $k$  at some intermediate time  $u$  (summed over all possible  $k$ 's)
3. to state  $j$  at time  $t$

gives us the following equation:

$$P_{ij}(v, t) = \sum_{k=1}^{\infty} P_{ik}(v, u) P_{kj}(u, t)$$

#### 2.2.0.22 Partial of $P_{ij}(v, t)$

$$\frac{\partial P_{ij}(v, t)}{\partial t} = \lim_{\Delta t \rightarrow 0} \sum_{k \neq j} P_{ik}(v, t) \frac{P_{kj}(t, t + \Delta t) - P_{kj}(t, t)}{\Delta t} + P_{ij}(v, t) \frac{P_{jj}(t, t + \Delta t) - 1}{\Delta t}$$

For a derivation of this, see (3.2.0.28).

**2.2.0.23 Rate Definitions for Simplification** We define  $q_{ij}$  in order to simplify the partial expression above. Additionally, we define  $q_{ij}$  in such a way that  $\sum_j q_{ij}(t) = 0 \forall i, t$

$$q_{kj}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{kj}(t, t + \Delta t) - P_{kj}(t, t)}{\Delta t} \quad k \neq j$$

$$q_{jj}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{jj}(t, t + \Delta t) - 1}{\Delta t}$$

This gives us the intended simplification of the partial:

$$\frac{\partial P_{ij}(v, t)}{\partial t} = \lim_{\Delta t \rightarrow 0} \sum_{k \neq j} P_{ik}(v, t) q_{kj}(t) + P_{ij}(v, t) q_{jj}(t)$$

These  $q_{ij}$ 's become useful later on as they become the arrival and departure rates in a birth-death process.

**2.2.0.24 State Probability** Let  $\pi_j(t)$  be the probability of being in state  $j$ , at time  $t$ .

$$\pi_j(t) = P(X(t) = j)$$

By applying total probability, over some earlier time  $v$  we get:

$$\pi_j(t) = \sum_i P_{ij}(v, t) \pi_i(v) \quad (\text{which is identical to the discrete case})$$

Taking the derivative we arrive at (see derivation (3.2.0.29)):

$$\frac{d\pi_j(t)}{dt} = \sum_{k \neq j} \pi_k(t)q_{kj}(t) + \pi_j(t)q_{jj}(t)$$

Applying the assumption that the transition probabilities are homogeneous (independent over time), we have that  $q_{ij}(t) = q_{ij}$

$$\frac{d\pi_j(t)}{dt} = \sum_{k \neq j} \pi_k(t)q_{kj} + \pi_j(t)q_{jj}$$

#### 2.2.0.25 Steady State Results

#### 2.2.0.26 Residence Time in State

## 3 Derivations

### 3.1 Probability

### 3.2 Stochastic Processes

**3.2.0.27 Chapman-Kolmogorov Equation** For context see (2.1.0.10). TODO

**3.2.0.28 Partial of  $P_{ij}(v, t)$**  For context see (2.2.0.22). TODO

**3.2.0.29 Derivative of  $\pi_j(t)$**  For context see (2.2.0.24).