Lecture 1 Ch 1. Markov chains

1.1 Definitions and Examples

Example 1.1. Gambler’s ruin. On any turn you win $1 with probability $p = 0.4$ and lose $1$ with probability $0.6$. Let $X_n$ be your total gain after $n$ games. Then

$$P(X_{n+1} = i + 1 \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = 0.4$$

for any possible history $i_0, \ldots, i_{n-1}, i$. Recall that

$$P(A \mid B) = \ldots \text{ always } P(AB) = P(A \mid B)P(B),$$

$$P(A_n A_1) = P(A_n \mid A_{n-1} \ldots A_1) P(A_2 \mid A_1) P(A_1).$$

You decided to stop once you total gain reaches $N$. Casino stops you when you lose all the money you started with.

Definition. Given a sequence of random variables $X_n, n = 0, 1, 2, \ldots$, taking values in a set $S = \{s_1, s_2, \ldots\}$ and a real-valued function $p(i, j), i, j \in S$, we say that $X_n$ is a discrete time Markov chain with transition matrix (or function) $p$ if for any $n \geq 0, j_0, j_1, \ldots, j_{n+1} \in S$

$$P(X_{n+1} = j_{n+1} \mid X_n = j_n, \ldots, X_0 = j_0) = p(j_n, j_{n+1}),$$

provided ...

$S$ is the state space, the distribution of $X_0$ is the initial distribution.

How is the matrix $(p(i, j))$ represented as a table, rows, columns?

Temporally homogeneous

Properties of transition matrix

$$p(i, j) \geq 0, \quad \sum_j p(i, j) = 1.$$ 

Can start the chain from any state, roll a die or generate a random number on a computer...

Transition matrix for Gambler’s ruin if $N = 5$

Pictorial representation $X \rightarrow Y$ ...

Example 1.2. Ehrenfest chain. Two boxes of air connected by a small hole. Total $N$ molecules. We pick one at random and move it to another box. Let $X_n$ be the number of molecules in the left box at the $n$th step.

$$P(X_{n+1} = i + 1 \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \frac{N - i}{N},$$

$$p(i, i + 1) = \frac{N - i}{N}, \quad p(i, i - 1) = \frac{i}{N}.$$
$N = 4$ write down the matrix

Example 1.7. Repair chain. A machine has three critical parts that are subject to failure, but can function as long as two of three parts are working. When two are broken, they are replaced and the machine is back to working order the next day. We assume that parts 1,2,3 fail with probability 0.01, 0.02, 0.04 but no two parts fail on the same day.

State space

$$\{0, 1, 2, 3, 12, 13, 23\}$$

The transition matrix

$$p(1, 12) = 0.02, \quad p(1, 13) = 0.04, \quad p(1, 1) = 0.94...$$

$$p = \begin{pmatrix}
0.93 & 0.01 & 0.02 & 0.04 & 0 & 0 & 0 \\
0 & 0.94 & 0 & 0.02 & 0.04 & 0 \\
0 & 0 & 0.95 & 0 & 0.01 & 0.04 \\
0 & 0 & 0 & 0.97 & 0 & 0.01 & 0.02 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}$$

Question; If we are going to operate the machine for 1800 days ($\sim 5$ years), then how many parts of type 1,2,3 will we need?

1.2 Multistep Transition Probabilities. p. 9.

$$p^m(i, j) = P(X_{n+m} = j \mid X_n = i).$$

Theorem 1.1. p. 11. The matrix $p^m = (p^m(i, j))$ is (indeed) the $m$th power of $p = (p(i, j)) = p^1$.

Recall how one multiplies two matrices, three matrices.

Proof of Theorem 1.1.

$$P(X_{n+1} = j_{n+1}, X_n = j_n, ..., X_0 = j_0)$$

$$= p(j_n, j_{n+1})p(j_{n-1}, j_n) \cdot \cdot \cdot p(j_0, j_1)P(X_0 = j_0)$$

$$= P(X_0 = j_0)p(j_0, j_1) \cdot \cdot \cdot p(j_{n-1}, j_n)p(j_n, j_{n+1}).$$

It follows that, for $k \leq n$

$$P(X_{n+1} = j_{n+1}, X_n = j_n, ..., X_k = j_k)$$

$$= p(X_k = j_k)p(j_k, j_{k+1}) \cdot \cdot \cdot p(j_{n-1}, j_n)p(j_n, j_{n+1})$$

$$P(X_{n+1} = j_{n+1}, X_n = j_n, ..., X_{k+1} = j_{k+1} \mid X_k = j_k)$$

$$= p(j_k, j_{k+1}) \cdot \cdot \cdot p(j_{n-1}, j_n)p(j_n, j_{n+1}),$$

whenever... Summing with $j_n, ..., j_k$ we get that $p^{n+1-k} = p^{n+1-k}$ and setting $k = n$ we get $p^1 = p$. 

Lecture 2
Observation. Good to know that if \( n \geq 1 \) and a set \( A \subset S^n \)

\[
P((X_0, \ldots, X_{n-1}) \in A, X_n = j_n) = \sum_{(j_0, \ldots, j_{n-1}) \in A} P(X_0 = j_0, \ldots, X_n = j_n) p(j_n, j_{n+1})
\]

It follows that

\[
P(X_{n+1} = j_{n+1} \mid X_n = j_n, (X_0, \ldots, X_{n-1}) \in A) = p(j_n, j_{n+1})
\]

provided...

1.3. Classification of states.

Notation

\[
P_x(A) = P(A \mid X_0 = x)
\]

Assume the chain starts at \( y \)

\[
T_y = \min\{n \geq 1 : X_n = y\} \quad (\min \emptyset = \infty)
\]

is the time of the first return to \( y \).

\[
\rho_{yy} = P_y(T_y < \infty).
\]

Definition A \( \{\infty, 0, 1, 2, \ldots\} \)-valued random variable \( T \) is called a stopping time if for any \( n = 0, 1, \ldots, n < \infty \), there exists \( G_n \subset S^{n+1} \) such that

\[
\{T = n\} = \{(X_0, \ldots, X_n) \in G_n\}
\]

that is if the occurrence (or nonoccurrence) of the event \( \{T = n\} \) can be determined by observing the values of the chain \( X_0, \ldots, X_n \) up to time \( n \).

Example: \( T_y \).

Theorem 1.2. Strong Markov property. p. 13. Suppose \( T \) is a stopping time. Given that \( T = n \) and \( X_T = y \), any other information about \( X_0, \ldots, X_T \) is irrelevant for predicting the future, and \( Y_k := X_{T+k}, \ k \geq 0 \), is a Markov chain with initial state \( y \) and transition function \( p \).

Theorem Suppose \( T \) is a stopping time. Given that \( T < \infty \) the process \( Y_k := X_{T+k}, \ k \geq 0 \), is a Markov chain with transition function \( p \) and initial distribution which is the conditional distribution of \( X_T \) given that \( T < \infty \).

Proof (of the last theorem). We have to show that

\[
P(X_{T+n+1} = j_{n+1}, X_{T+n} = j_n, \ldots, X_T = j_0 \mid T < \infty) = p(j_n, j_{n+1}) P(X_{T+n} = j_n, \ldots, X_T = j_0 \mid T < \infty)
\]

May assume \( P(T < \infty) > 0 \). Then we need

\[
P(X_{T+n+1} = j_{n+1}, X_{T+n} = j_n, \ldots, X_T = j_0, T < \infty) = p(j_n, j_{n+1}) P(X_{T+n} = j_n, \ldots, X_T = j_0, T < \infty).
\]
The left-hand side is
\[
\sum_{k=0}^{\infty} P(X_{T+n+1} = j_{n+1}, X_{T+n} = j_n, \ldots, X_T = j_0, T = k)
\]
\[
= \sum_{k=0}^{\infty} P(X_{k+n+1} = j_{n+1}, X_{k+n} = j_n, \ldots, X_k = j_0, (X_0, \ldots, X_k) \in G_k)
\]
\[
= p(j_n, j_{n+1}) \sum_{k=0}^{\infty} P(X_{T+n} = j_n, \ldots, X_T = j_0, T = k)
\]
\[
= p(j_n, j_{n+1}) \sum_{k=0}^{\infty} P(X_{T+n} = j_n, \ldots, X_T = j_0, T < \infty).
\]

\[\square\]

**Corollary** For any \( k \geq 0 \) and \( G_k \subset S^{k+1} \)
\[
P((X_T, \ldots, X_{T+k}) \in G_k \mid T < \infty) = P_x(X_0, \ldots, X_k) \in G_k),
\]
provided \( P(T < \infty) > 0 \), where \( X_n \) is the Markov chain with transition function \( p \) and initial distribution \( \pi(i) = P(X_T = i \mid T < \infty) \).

What is the probability of two returns? If \( T_y < \infty \) define
\[
T_y^{(2)} = \min\{k \geq 1 : X_{T_y+k} = y\}.
\]

We have
\[
P_y(T_y^{(2)} < \infty) = P_y(T_y^{(2)} < \infty \mid T_y < \infty) \rho_{yy}
\]
\[
= P_y(\exists k : X_{T_y+k} = y \mid T_y < \infty) \rho_{yy}
\]
\[
= P_y(\min\{k \geq 1 : X_k = y\} < \infty) g_y = \rho_{yy}^2.
\]

Let \( T_y^1 = T_y \),
\[
T_y^k = \min\{n \geq T_y^{k-1} + 1 : X_n = y\},
\]
the time of the \( k \)th return to \( y \).
\[
P_y(T_y^k < \infty) = \rho_{yy}^k
\]

**Definition** If \( \rho_{yy} < 1 \), then \( y \) is transient. If \( \rho_{yy} = 1 \), then \( y \) is recurrent.

**Theorem** Assume the chain starts at \( y \). If \( y \) is recurrent, then \( V(y) = \infty \) (a.s.). If \( y \) is transient, \( V(y) \) has a geometric distribution with parameter \( \rho_{yy} \) and mean \( 1/(1 - \rho_{yy}) \).

**Proof.** We have for \( n \geq 0 \) (\( T_y^n := 0 \))
\[
P_y(V(y) > n) = P_y(T_y^n < \infty) = \rho_{yy}^n.
\]
\[\square\]
Corollary. Theorem 1.13. \( y \) is recurrent iff
\[
\sum_{n=1}^{\infty} p^n(y, y) = \infty.
\]
Indeed
\[
V(y) = \sum_{n=0}^{\infty} I_y(X_n)
\]
and
\[
\frac{1}{1 - \rho_{yy}} = E_y V(y) = 1 + \sum_{n=1}^{\infty} p^n(y, y).
\]
\[\square\]

Example 1.12. Gambler’s ruin. \( N = 4, P(\text{win}) = 0.4 \), transient and recurrent states...

Definition 1.1. We say that \( x \) communicates with \( y \) and write \( x \to y \) if
\[
\rho_{xy} := P_x(T_y < \infty) > 0
\]
that is if there exists an \( n \) and a path \( x_0, ..., x_n \) such that \( x_0 = x, x_n = y \) and \( p(x_i, x_{i+1}) > 0 \) for \( i = 0, ..., n - 1 \).

Lemma 1.4. If \( x \to y \) and \( y \to z \), then \( x \to z \).

Theorem 1.5. p. 16. If \( \rho_{xy} > 0 \) but \( \rho_{yx} < 1 \), then \( x \) is transient.

Proof. We have
\[
P_x(T_x = \infty) \geq \rho_{xy}(1 - \rho_{yx}).
\]
\[\square\]

~ Lemma 1.6+1.9. If \( x \) is recurrent and \( \rho_{xy} > 0 \), then \( \rho_{yx} = 1 \), \( y \) is recurrent, and \( \rho_{xy} = 1 \).

Proof. Assuming \( \rho_{yx} < 1 \) leads to contradiction with Theorem 1.5. Each time the chain visits \( x \) there is a nonzero probability to visit \( y \). The chain started at \( x \) visits \( x \) infinitely many times, therefore, it will visit \( y \) infinitely many times, so \( y \) is recurrent. Since the chain started at \( x \) visits \( y \) with probability one infinitely often, it does visit \( y \) with probability one: \( \rho_{xy} = 1 \). \[\square\]

Lecture 3

Example 1.14 (A seven-state chain, p. 17).
\[
P = \begin{pmatrix}
0.7 & 0 & 0 & 0 & 0.3 & 0 & 0 \\
0.1 & 0.2 & 0.3 & 0.4 & 0 & 0 & 0 \\
0 & 0.5 & 0.3 & 0.2 & 0 & 0 & 0 \\
0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 \\
0.6 & 0 & 0 & 0 & 0.4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.2 & 0.8 \\
0 & 0 & 0 & 0.1 & 0 & 0 & 0 \\
\end{pmatrix}
\]
The graph is

1 ← 2 → 4 → 6
↑↓ ↓↗ ↑↙
5 ← 3 7

2,3 are transient. All the remaining ones are recurrent.

Definition. A set $A$ is closed if it is impossible to get out: $i \in A, j \notin A$ implies $\rho_{ij} = 0$.

Closed sets in Example 1.14...

Definition. A set $B$ is irreducible if, whenever $i,j \in B$, $i$ communicates with $j$.

Irreducible sets in Example 1.14...

Example. Weather chain.

$$p = \begin{pmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{pmatrix}$$

$P_1(T_1 > n) \leq (0.9)^n$. All states are recurrent.

Lemma 1.3. (Pedestrian lemma. p. 16.) Let $k$ be such that $P_x(T_y \leq k) \geq \alpha > 0$ for all $x$. Then

$$P_x(T_y > nk) \leq (1 - \alpha)^n.$$ 

Proof. We have

$$P_x(T_y > nk + k) = P_x(T_y > nk + k, T_y > nk)$$

$$= \sum_{j \neq y} P_x(T_y > nk + k, X_{nk} = j, T_y > nk)$$

$$= \sum_{j \neq y} P_x(X_{nk+1} \neq y, ..., X_{nk+k} \neq y | X_{nk} = j)P(X_{nk} = j, T_y > nk)$$

$$= \sum_{j \neq y} P_j(T_y \geq k)P(X_{nk} = j, T_y > nk)$$

$$\leq (1 - \alpha) \sum_{j \neq y} P(X_{nk} = j, T_y > nk)$$

$$\leq (1 - \alpha)P(T_y > nk).$$

□

Theorem 1.7. If $C$ is a finite, closed, and irreducible set, then all states in $C$ are recurrent.

Proof. Let $N$ be such that $P_x(T_y \leq N) \geq \alpha > 0$ for any $x,y \in C$. Then apply Lemma 1.3. □
Theorem 1.8. If the state space $S$ is finite, then $S$ can be written as a disjoint union $T \cup R_1 \cup \ldots \cup R_k$, where $T$ is a set of transient states and the $R_i$, $1 \leq i \leq k$, are closed irreducible sets of recurrent states.

Proof. Let $T = \{x : \exists y : x \rightarrow y, y \not\rightarrow x\}$. The states in $T$ are transient.

Take $x \in S \setminus T$. Let $R_x$ be the set of $y$ with which $x$ communicates. Each $y \in R(x)$ communicates with $x$ and then each $y \in R(x)$ communicates with each $z \in R(x)$ and does not communicate with outside states. Then $R(x)$ is closed and irreducible. If $(S \setminus T) \setminus R(x) \neq \emptyset$ repeat the procedure. □

1.4. Stationary distributions. p.21

If the initial distribution is $q(i) = P(X_0 = i)$

$$P(X_n = j) = \sum_i q(i)p^n(i,j).$$

Row vector, matrix multiplication.

If $q = qp$ $q$ is a stationary distribution. Use $\pi$ for solutions of

$$\pi = \pi p$$

such that $\pi(i) \geq 0$ and $\sum_i \pi(i) = 1$. Chain started from a stationary distribution... $\pi = \pi p^\infty$...There could be very many stationary distribution...

Remark If we have $q$ such that $q = qp$ and $Q := \sum |q(i)| < \infty$, then $\pi(i) := |q(i)|/Q$ is a stationary distribution.

Lemma * If the chain is irreducible and a stationary distribution exists, then $\pi(x) > 0$ for any $x$.

Proof. Find $y$ such that $\pi(y) > 0$ and find $n$ such that $p^n(y,x) > 0$. □

Lemma 1.25, p. 50 If there is a stationary distribution, then all states $y$ for which $\pi(y) > 0$ are recurrent.

Proof. Let $N(y)$ denote the number of visits to $y$ at times $n \geq 1$. We have

$$E_x N(y) = \sum_{n=1}^{\infty} p^n(x,y),$$

$$\sum_x \pi(x)E_x N(y) = \sum_n \sum_x \pi(x)p^n(x,y) = \sum_n \pi(y) = \infty,$$

and

$$E_x N(y) \leq E_y N(y) = \frac{1}{1 - \rho_{yy}}, \quad \sum_x \pi(x)E_x N(y) \leq \frac{1}{1 - \rho_{yy}}.$$ 

Hence $\rho_{yy} = 1$. □

Lecture 4

Corollary* If the chain is irreducible and a stationary distribution exists, then all states are recurrent and $\rho_{xy} = 1$ for all $x, y \in S$. General two state transition probability.
\[ p = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix}. \]

\[ \pi_1 = \frac{b}{a + b}, \quad \pi_2 = \frac{a}{a + b}. \]

**Theorem** If \( S \) is finite, then stationary distributions exist (could be many).

Proof. Fix \( i \in S \). The sequences

\[ \frac{1}{n} \sum_{k=1}^{n} p^k(i,j) \]

are bounded. There exist \( n_m \to \infty \) such that for any \( j \in S \) the limit

\[ \pi(j) := \lim_{m \to \infty} \frac{1}{n_m} \sum_{k=1}^{n_m} p^k(i,j) \]

exists. \( \pi(j) \geq 0, \sum_j \pi(j) = 1, \)

\[ \sum_j \pi(j)p(j,r) = \lim_{m \to \infty} \frac{1}{n_m} \sum_{k=1}^{n_m} p^{k+1}(i,r) \]

\[ = \lim_{m \to \infty} \frac{1}{n_m} \left( \sum_{k=1}^{n_m} p^k(i,r) - p(i,r) + p^{n+1}(i,r) \right). \]

\[ \square \]

**Fact** If \( S \) is finite and irreducible, then the stationary distribution \( \pi \) is unique and (the law of large numbers) for any \( x \in S \)

\[ \lim_{n \to \infty} \frac{N_n(x)}{n} = \pi(x). \]

**Example 1.18. Social mobility. Cont of Ex. 1.4. p. 23.** Let \( X_n \) be a family’s social class in the \( n \)th generation, which we assume either 1=lower, 2=middle, or 3=upper. Assume this is a Markov chain with

\[ p = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}. \]

System:

Give a man a fish and you feed him for a day. Teach a man to fish and you feed him for a lifetime.

\( \pi_1 = 22/47, \pi_2 = 16/47, \pi_3 = 9/47. \)

**Doubly Stochastic Chains**

**Definition...**

**Theorem 1.14. p. 26.** If \( p \) is doubly stochastic and \( \#S = N \), then \( \pi(i) = 1/N, i \in S \), is a stationary distribution.
Symmetric reflective random walk on the line $S = \{0, 1, ..., L\}$. If $L = 4$

$$p = \begin{pmatrix}
    1/2 & 1/2 & 0 & 0 & 0 \\
    1/2 & 0 & 1/2 & 0 & 0 \\
    0 & 1/2 & 0 & 1/2 & 0 \\
    0 & 0 & 1/2 & 0 & 1/2 \\
    0 & 0 & 0 & 1/2 & 1/2
\end{pmatrix}$$


Detailed balance condition

$$\pi(x)p(x, y) = \pi(y)p(y, x)$$

If $\pi \geq 0$ and $\sum \pi = 1$, $\pi$ is automatically stationary.

Example 1.26. Birth and death chains. State space $S = \{l, l+1, ..., r-1, r\}$ integers and $p(x, y) = 0$, if $|x - y| > 1$. Let

- $p(x, x + 1) = p_x$ for $x < r$,
- $p(x, x - 1) = q_x$ for $x > l$,
- $p(x, x) = 1 - p_x - q_x$ for $l < x < r$,

$p(r, r) = 1 - q_r$, $p(l, l) = 1 - p_l$. All other $p$ are zero.

If $x < r$, detailed balance: $\pi(x)p_x = \pi(x + 1)q_{x+1}$,

$$\pi(x + 1) = \frac{p_x}{q_{x+1}}\pi(x),$$

$$\pi(l + i) = \pi(l)\frac{p_{l+i-1}p_{l+i-2} \cdots p_{l+1}p_l}{q_{l+i}q_{l+i-1} \cdots q_{l+2}q_{l+1}}.$$ 

Example 1.27. Ehrenfest chain. p. 30. Total number of balls is $N$. $X_k$ is the number of balls in the left box at the $k$th step.

$$p(i, i + 1) = \frac{N - i}{N} \quad i < N, \quad p(i, i - 1) = \frac{i}{N} \quad i \geq 0.$$ 

Stationary distribution:

$$\pi(i)p(i, i + 1) = \pi(i + 1)p(i + 1, i),$$

$$\pi(i + 1) = \frac{N - i}{i + 1}\pi(i) = \frac{(N - i)(N - i + 1) \cdots N}{(i + 1) \cdots 1}\pi(0) = \left(\frac{N}{i + 1}\right)\pi(0).$$

$$\pi(i) = 2^{-N}\binom{N}{i}.$$

Example 1.29. Random walks on graphs. A graph is described by giving two things: (i) a set of vertices $V$ (a finite set) and (ii) an adjacency matrix
with entries $A(u, v)$ which is 1 if there is an edge connecting $u$ and $v$ and zero otherwise. By convention we set $A(v, v) = 0$. The degree of $u$:

$$d(u) = \sum_v A(u, v).$$

Transition probability function

$$p(u, v) = \frac{A(u, v)}{d(u)}, \quad A(u, v) = A(v, u).$$

Stationary distribution $\pi(u) = cd(u), c = \frac{1}{\sum_u d(u)}$.

Example. Random walk on a 4 checkerboard

The degrees

$$\begin{pmatrix}
2 & 3 & 3 & 2 \\
3 & 4 & 4 & 3 \\
3 & 4 & 4 & 3 \\
2 & 3 & 3 & 2
\end{pmatrix}$$

The sum of degrees = 48.

Extra Example. Symmetric random walk.

$$p^{2n}(0, 0) = \left(\frac{2n}{n}\right)^{2-2n}, \quad n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

$$\frac{(2n)!}{(n!)^2} \sim \frac{\sqrt{2(2n/e)^{2n}}}{\sqrt{2\pi n(n/e)^{2n}}}, \quad p^{2n}(0, 0) \sim \frac{1}{\sqrt{\pi n}}, \quad \sum_n p^{2n}(0, 0) = \infty,$$

0 is recurrent.

Lecture 5

Above examples were dealing with finite-state chains. Example* $S = \{0, 1, 2, ...\}, p(n, n + 1) = 1 - 1/(n + 2)^2, p(n, 0) = 1/(n + 2)^2$

$$P_0(T_0 = \infty) = \prod_{n=0}^{\infty} \left(1 - \frac{1}{(n + 2)^2}\right) > 0.$$ No stationary distribution.

1.6 Limit behavior. p. 40

If $y$ is transient $\sum p^n(x, y) < \infty$ and $p^n(x, y) \to 0$ as $n \to \infty$.

Let $S$ be finite, still $\lim p^n(x, y)$ may not exist.

Example 1.36. Ehrenfest Chain from 1.2. $N = 3, S = \{0, 1, 2, 3\}$,

$$p = \begin{pmatrix}
0 & 1 & 0 & 0 \\
1/3 & 0 & 2/3 & 0 \\
0 & 2/3 & 0 & 1/3 \\
0 & 0 & 1 & 0
\end{pmatrix}, \quad p^2 = \begin{pmatrix}
1/3 & 0 & 2/3 & 0 \\
0 & 7/9 & 0 & 2/9 \\
2/9 & 0 & 7/9 & 0 \\
0 & 2/3 & 0 & 1/3
\end{pmatrix}.$$  

Definition The period of a state $x$ is the largest number that divides all the $n \geq 1$ for which $p^n(x, x) > 0$. That is it is the greatest common divisor of $I_x = \{n \geq 1 : p^n(x, x) > 0\}$. 

Observation. Lemma 1.27. $I_x$ is closed under addition.

**Lemma 1.17** If $x$ and $y$ communicate with each other, they have the same period.

**Proof.** Let the period of $x$ be $c$ and that of $y$ be $d$. Let $k, m$ be such that $p^k(x, y) > 0$ and $p^m(y, x) > 0$. Then

$$k + m \in I_x, \quad (k + m)/c \in \{1, 2, \ldots\}$$

and for any $l \in I_y$ we have $k + m + l \in I_x$. Then $l/c$ is an integer, but then by the definition of the greatest common divisor we have $c \leq d$. Interchanging $x$ and $y$ we get $d \leq c$. \(\square\)

Compare with Example*.

**Lemma 1.18** If $p(x, x) > 0$, then $x$ has period 1.

**Lemma** Let $I$ be a set of integers $\geq 1$ and $d$ be their greatest common divisor. Then there exist $k \geq 1, a_1, \ldots, a_k \in I$ and $i_1, \ldots, i_k \in \{\pm 1, \pm 2, \ldots\}$ such that

$$i_1a_1 + \ldots + i_ka_k = d.$$ 

**Proof.** Take all linear combinations that are $> 0$ and take the one for which

$$i_1a_1 + \ldots + i_ka_k = d_{\text{min}}$$

is minimal. $d_{\text{min}}$ is divisible by $d$, hence $d_{\text{min}} \geq d$.

**Claim:** $I \subset \{d_{\text{min}}, 2d_{\text{min}}, \ldots\}$. Indeed, assume $a_{k+1} \in I$ and $rd_{\text{min}} < a_{k+1} < (r + 1)d_{\text{min}}$. Then

$$0 < a_{k+1} - r(i_1a_1 + \ldots + i_ka_k) < d_{\text{min}},$$

which is a contradiction.

The claim implies that $d_{\text{min}}$ divides all elements of $I$ and hence $d \geq d_{\text{min}}$. \(\square\)

**Lemma 1.26.** p. 51. If $x$ has period 1, then there is a number $n_0$ such that $p^n(x, x) > 0$ for all $n \geq n_0$.

**Proof.** Let $i_1, \ldots, i_m > 0, i_{m+1}, \ldots, i_k < 0$ and

$$i_1a_1 + \ldots + i_ma_m = 1 + |i_{m+1}|a_{m+1} + \ldots + |i_k|a_k =: J.$$ 

Then $J \in I_x, J^2 \in I_x$, and $J + 1 \in I_x$ and, for any integers $p \geq 0$ and $0 \leq q \leq J$,

$$J^2 + pJ + q = J(J - q) + pJ + (J + 1)q \in I_x.$$ 

Hence $n_0 = J^2$ does the job. \(\square\)

**Theorems 1.19, 1.23** Suppose $p$ is irreducible, aperiodic, and has stationary distribution $\pi$. Then $p^n(x, y) \to \pi(y)$ for any $x, y \in S$ as $n \to \infty$. Furthermore

$$\sum_y |p^n(x, y) - \pi(y)| \to 0.$$


Proof. By coupling. Define
\[ S^2 = S \times S, \quad \bar{p}(x_1, y_1, x_2, y_2) = p(x_1, x_2)p(y_1, y_2). \]
Since both chains are irreducible and aperiodic, \((X_n, Y_n)\) is irreducible (needs explanation based on aperiodicity).
\[ \bar{\pi}(a, b) = \pi(a)\pi(b) \] is stationary.
Start \(X_n\) with starting point \(x\) and \(Y_0\) having the initial distribution \(\pi\).
Corollary* all states in \(S^2\) are recurrent, in particular,
\[ V_{(x,x)} = \min(n \geq 0 : X_n = Y_n = x) < \infty. \]
Define
\[ T = \min(n \geq 0 : X_n = Y_n), \quad T \leq V_{(x,x)} \implies T < \infty. \]
On \(\{T \leq n\}\) the two components \(X_n\) and \(Y_n\) have the same distribution:
\[ P(X_n = y, T \leq n) = \sum_{m=0}^{n} \sum_{x} P(X_n = y, X_m = x, T = m) \]
\[ = \sum_{m=0}^{n} \sum_{x} P(X_n = y | X_m = x)P(X_m = y, T = m) \]
\[ = \sum_{m=0}^{n} \sum_{x} P(Y_n = y | Y_m = x)P(Y_m = y, T = m) = P(Y_n = y, T \leq n). \]
Since
\[ P(X_n = y) = P(X_n = y, T \leq n) + P(X_n = y, T > n) \]
\[ = P(Y_n = y, T \leq n) + P(X_n = y, T > n), \]
and
\[ P(Y_n = y) = P(Y_n = y, T \leq n) + P(Y_n = y, T > n), \]
we have
\[ \sum_{y} |P(X_n = y) - P(Y_n = y)| \]
\[ \leq \sum_{y} (P(X_n = y, T > n) + P(Y_n = y, T > n)) = 2P(T > n). \]
Here \(P(X_n = y) = p^n(x, y), P(Y_n = y) = \pi(y). \]

Lecture 6
1.7 Returns to a fixed state. p. 46

Theorem 1.20 (Law of large numbers.) Suppose the chain is irreducible and reaches \(y\) with probability one. Let \(N_n(y)\) be the number of visits to \(y\) at times \(\leq n\). Then as \(n \to \infty\)
\[ \frac{N_n(y)}{n} \to \frac{1}{E_y T_y}. \]

Proof. Case 1. State \(y\) is transient. Then the number of visits to \(y\) is finite with probability one and \(E_y T_y = \infty\).
Case 2. State \( y \) is recurrent. Define \( R_n \) to be the number of visits to \( y \) after the first visit there. Let \( t(0), t(1), \ldots \) be the times of the first, the second, ... visits to \( y \). Then \( t(n + 1) - t(n) \) are iid and by the strong law of large numbers
\[
\frac{t(n)}{n} \to E(t_1 - t_0) = E_y T_y.
\]
It follows that
\[
\frac{t(N_n(y) + 1)}{N_n(y)}, \quad \frac{t(N_n(y))}{N_n(y)} \to E(t_1 - t_0) = E_y T_y.
\]
However, \( t(N_n(y)) \leq n \leq t(N_n(y) + 1) \). \( \square \)

THEOREM 1.21. If the chain is irreducible and has stationary distribution, then (all states are recurrent, \( \rho_{xy} = 1 \), and)
\[
\pi(y) = \frac{1}{E_y T_y}.
\]
Proof. Start the chain from the stationary distribution, note that \( E_{\pi} N_n(y) = n\pi(y) \) and use the dominated convergence theorem.

Corollary of this and Lemma * If the chain is irreducible and has stationary distribution, then all states are positive recurrent: \( E_y T_y < \infty \) for all \( y \) and
\[
\frac{N_n(y)}{n} \to \pi(y).
\]

Corollary. If the chain is irreducible, then it can have only one stationary distribution. May have none.

Null recurrent...Compare with Example*. No stationary distribution.

Theorem 1.22. Suppose \( p \) is irreducible, has stationary distribution \( \pi \), and \( \sum_x |f(x)|\pi(x) < \infty \). Then, as \( n \to \infty \),
\[
\frac{1}{n} \sum_{m=1}^{n} f(X_m) \to \sum_x f(x)\pi(x).
\]
Proof for \( f \) bounded. It suffices to prove this for \( 1 \geq f \geq 0 \), in which case we observe that
\[
I_n := \frac{1}{n} \sum_{m=1}^{n} f(X_m) = \sum_{y \in S} f(y) \frac{N_n(y)}{n}.
\]
Next, for any \( \varepsilon > 0 \) we can find a finite subset \( S' \) of \( S \) such that
\[
\sum_{y \in S'} f(y)\pi(y) \geq \sum_{y \in S} f(y)\pi(y) - \varepsilon
\]
and for any trajectory of the chain we can find \( n_0 \) such that for \( n \geq n_0 \)
\[
\sum_{y \in S'} f(y) \frac{N_n(y)}{n} \geq \sum_{y \in S'} f(y)\pi(y) - \varepsilon,
\]
implying that for \( n \geq n_0 \) we have

\[
I_n \geq \sum_{y \in S} f(y)\pi(y) - 2\varepsilon.
\]

By substituting here \( 1 - f \) in place of \( f \) we find that perhaps for a larger \( n_0 \) and \( n \geq n_0 \) we also have

\[
1 - I_n \geq \sum_{y \in S} (1 - f)(y)\pi(y) - 2\varepsilon = 1 - \sum_{y \in S} f(y)\pi(y) - 2\varepsilon
\]

and the result follows. \( \square \)

How about stationary measures?

Theorem 1.24, p. 48. Suppose \( p \) is irreducible and recurrent. Take \( x \in S \) and let \( T_x = \inf\{n \geq 1, X_n = x\} \). Then

\[
\mu_x(y) = \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n)
\]

defines a stationary measure with \( 0 < \mu_x(y) < \infty \).

Proof. Case 1. \( z \neq x \). We have

\[
\sum_y \mu_x(y)p(y,z) = \sum_y \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n)p(y,z)
\]

\[
= \sum_{n=0}^{\infty} \sum_y P_x(X_n = y, X_{n+1} = z, T_x > n+1)
\]

\[
= \sum_{n=0}^{\infty} P_x(X_{n+1} = z, T_x > n+1).
\]

Case 2. \( z = x \). We have

\[
\sum_y \mu_x(y)p(y,x) = \sum_{n=0}^{\infty} \sum_y P_x(X_n = y, X_{n+1} = x, T_x = n+1)
\]

\[
= P_x(T_x < \infty) = 1 = P_x(X_0 = x, T_x > 0) = \mu_x(x).
\]

That \( \mu > 0 \) follows from the fact that the chain is irreducible and you can reach \( y \) starting from \( x \) before coming back to \( x \) with positive probability. To prove that \( \mu_x(y) < \infty \) \( (y \neq x, \text{note that } \mu(x) = 1) \), observe that

\[
\mu_x(y) = E_x \sum_{n=0}^{T_x-1} I_{X_n = y},
\]

where

\[
\sum_{n=0}^{T_x-1} I_{X_n = y}
\]

is the number of visits to \( y \) while never visiting \( x \). This number has a geometric distribution with parameter less than 1, because the chain is irreducible and the probability to return to \( y \) without hitting \( x \) is less than
one. Its mean is finite. A formal a posteriori proof is as follows: find \( n \) such that \( p^n(y,x) > 0 \) and note that

\[
1 = \mu_x(x) = \sum_z \mu_z(z)p^n(z,x) \geq \mu_x(y)p^n(y,x).
\]

**Observation** If \( x \) is positive recurrent, the above \( \mu/|\mu| \), where \( |\mu| = \sum \mu(y) = E_xT_x \), is a unique stationary distribution. One more time: If the chain is irreducible and recurrent and a stationary distribution exists, then all states are positive recurrent and the stationary distribution is unique.

Go to Example* \[
\mu_0(m) = P(X_m = m) = \prod_{i=0}^{m-1} \left( 1 - \frac{1}{(i + 2)^2} \right).
\]

**Lecture 7**

Theorem 1.30 p. 70 For irreducible chains the following are equivalent:
(i) Some \( x \) is positive recurrent;
(ii) There is a stationary distribution;
(iii) All states are positive recurrent.

1.9 Exit Distributions, p. 53

Example 1.41 (Tennis) In tennis the winner of a game is the first player to win four points unless the score is 4-3, in which case the game must continue until one player is ahead by two points and wins the game. Suppose that the server win the point with probability 0.6 and successive points are independent. What is the probability the server will win the game if the score is tied 3-3? if he is ahead by one point? behind by one point?

The state space is \( 2, 1, 0, -1, -2 \) with 2 win for server and -2 win for opponent. The states are difference in scores. Transition matrix

\[
p = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0.6 & 0 & 0.4 & 0 & 0 \\
0 & 0.6 & 0 & 0.4 & 0 \\
0 & 0 & 0.6 & 0 & 0.4 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\]

Let \( h(x) \) be the probability to win if the score is \( x \). Then

\[
h(x) = \sum_y p(x,y)h(y)
\]

with \( h(2) = 1 \) and \( h(-2) = 0 \). Write down the system. Solution:

\[(1, 0.8769, 0.6923, 0.4154, 0).\]

Given a set \( F \subset S \) let

\[
V_F = \min \{ n \geq 0 : X_n \in F \}.
\]
Lemma** Let $C$ be finite and $w(x)$ be a function such that $w \geq 0$ on $S \setminus C$ and for each $x \in C$
\[ w(x) \geq \sum_y p(x, y)w(y). \]

Also assume that $P_y(V_{S \setminus C} < \infty) > 0$ for any $y$. Then $w \geq 0$. In particular, if $w = 0$ on $S \setminus C$ and for each $x \in C$
\[ w(x) = \sum_y p(x, y)w(y), \]
then $w = 0$.

Proof. Let $x_0 \in C$ be such that $w(x_0) = \min_C w =: m$. Assume that $w(x_0) < 0$ and define $C' = \{ y \in C : w(y) = m \}$. It follows from
\[
m = w(x_0) = \sum_y p(x_0, y)w(y) = \sum_{y \in C} p(x_0, y)w(y)
= P_{x_0}(X_1 \in C')m + \sum_{y \in S \setminus C'} p(x_0, y)w(y) \geq m
\]
that $P_{x_0}(X_1 \in C') = 1$. Then we can take any other $y \in C'$ and get that $P_y(X_1 \in C') = 1$. Hence the chain never leaves $C'$ in contradiction with $P_y(V_{S \setminus C} < \infty) > 0$. Hence $\min_C w \geq 0$.

\[ \Box \]

Theorem 1.28 Consider a Markov chain with state space $S$. Let $A$ and $B$ be disjoint subsets of $S$, so that $C := S \setminus (A \cup B)$ is finite. If $P_x(V_A \wedge V_B < \infty) > 0$ for all $x \in C$, then $h(x) = P_x(V_A < V_B)$ is a unique function such that $h(x) = 1$ on $A$ and $h(x) = 0$ on $B$ and on $C$
\[ h(x) = \sum_y p(x, y)h(y). \] (1)

Proof. By the Markov property $h(x) = P_x(V_A < V_B)$ satisfies (1). It also equals 1 on $A$ and zero on $B$. Its uniqueness follows from Lemma**. \[ \Box \]

Example 1.41, Gambler’s Ruin. p. 57 Consider a gambling game in which on any turn you win $\$1 with probability $p$ and lose $\$1 with probability $q = 1 - p$. Suppose further that you will quit playing if your fortune reaches $\$N$. Of course, if your fortune reaches $\$0$, then the casino makes you stop. Let
\[ h(x) = P_x(V_N < V_0), \quad h(N) = 1, \quad h(0) = 0, \]
\[ h(x) = ph(x + 1) + qh(x - 1) \quad x \notin (0, N). \]
This is a finite-difference equation and we need to find its solutions. Try $\theta^x$
\[ 1 = p\theta + q\theta^{-1}, \quad \theta = 1 \quad \text{or} \quad \theta = q/p. \]

Remark* If $\theta \neq 1$,
\[ h(y) = c_1\theta^y + c_2 \]
is a nontrivial solution of $h(x) = ph(x + 1) + qh(x - 1)$ for $x \notin (0, N)$. 

\[ \Box \]
Continue the example. Let $\theta \neq 1$ and find $c_1, c_2$ so that
\[ c_1 \theta^0 + c_2 = 0, \quad c_1 \theta^N + c_2 = 1. \]

This yields
\[ h(y) = \frac{1 - \theta^y}{1 - \theta^N}. \]

Case $\theta = 1$. $h(y) = cy$, $cN = 1$, $c = 1/N$, $N \to \infty$.

Discussion $N \to \infty$. If $p > q$ the probability to get ruined tends to $\theta^x \neq 1$. It is one if $p \leq 1/2$.

1.10 Exit times, p. 61

Theorem 1.29, p. 62. Let $V_A = \inf\{n \geq 0 : X_n \in A\}$. Suppose that $C = S \setminus A$ is finite and $P_x(V_A < \infty) = 1$ for any $x \in C$. Then $g(x) = E_x V_A$ is a unique function such that $g = 0$ on $A$ and for $x \in C$
\[ g(x) = 1 + \sum_y p(x, y) g(y). \]

Proof. Lemma 1.3 implies that $g(x) < \infty$. The Markov property implies that $g(x) = E_x V_A$ satisfies the equation. Obviously, it is zero on $A$. Its uniqueness follows from Lemma**.

Lecture 8.

Example 1.52. Duration of fair games. Let $\tau = \min\{n \geq 0 : X_n \notin (0, N)\}$. $g(x) = E_x \tau$ satisfies
\[ g(x) = 1 + (1/2)g(x + 1) + (1/2)g(x - 1), \quad x \in (0, N), \]
\[ g(0) = g(N) = 0. \] The result $g(x) = x(N - x)$.

Example 1.52. Duration of nonfair games
\[ g(x) = 1 + pg(x + 1) + qg(x - 1), \quad x \in (0, N). \]

Observe that $h(x) = ax$ satisfies
\[ h(x) = ph(x + 1) + qh(x - 1) + a(q - p). \]

Hence $w = g - x/(q - p)$ satisfies
\[ w(x) = pw(x + 1) + qw(x - 1), \quad x \in (0, N). \]

By Remark*
\[ w(y) = c_1 \theta^y + c_2, \quad \theta = q/p, \]
also satisfies this equation. By Theorem 1.29
\[ E_x \tau = \frac{x}{q - p} + c_1 \theta^x + c_2 \]
if the constants $c_i$ are such that
\[ c_1 + c_2 = 0, \quad \frac{N}{q - p} + c_1 \theta^N + c_2 = 0. \]
The result is
\[ E_x \tau = \frac{x}{q - p} - \frac{N}{q - p} \frac{1 - \theta^x}{1 - \theta N}. \]

**Discussion** \( N \to \infty \). If \( q > p \), \( E_x \tau \to x/(q - p) \). If \( p \leq 1/2 \), \( E_x \tau \to \infty \). the point 0 is null recurrent for simple random walk. Interpretation with hockey games.

1.11. Infinite state space. Reflecting random walk. \( p(i, i + 1) = p \), \( i \geq 0 \), \( p(i, i - 1) = 1 - p \), \( i > 0 \), \( p(0, 0) = 1 - p \). This is a birth and death chain so we try to find the stationary distribution from the detailed balance condition
\[ p\pi(i) = (1 - p)\pi(i + 1), \quad i > 0. \]
\[ \pi(0) = c, \]
\[ \pi(i) = c(p/(1 - p))^i. \]
Case \( p < 1/2 \), \( \theta = p/q \), \( \pi(i) = (1 - \theta)^i \).
Aperiodic, irreducible, with stationary distribution,
\[ P(X_n = i) \to \pi(i) \text{ as } n \to \infty, \]
\[ E_0 T_0 = \frac{1}{\pi(0)} = \frac{1}{1 - \theta}. \]

Case \( p > 1/2 \), \( P_x(T_0 < \infty) = (1 - p)/p < 1 \), \( x > 0 \), all states are transient.

Case \( p = 1/2 \), \( P_0(T_0 < \infty) = 1 \) but \( E_0 T_0 = \infty \).

Example 1.55. Branching process. p. 70. Each individual in the \( n \)th generation gives birth to an independent and identically distributed number of children. \( P(\text{number of children} = k) = p_k \), \( k = 0, 1, 2, \ldots \)

What is the probability the species avoids extinction if it starts from one member?

**Lemma 1.31.** The extinction probability \( p \) is the smallest solution of the equation \( \phi(x) = x \), \( x \in [0, 1] \), where
\[ \phi(x) = \sum_{k=0}^{\infty} x^k p_k \]

Proof. Let \( \rho_n = P(X_n = 0) \). Then
\[ \rho_n = \sum_{k=0}^{\infty} p_k \rho_{n-1}^k, \]
\( \rho_n \) increase, \( \phi \) is continuous, \( \rho_n \to \rho \), where \( \rho \), the extinction probability, satisfies the equation.

If there is another solution \( \bar{\rho} \), we have \( 0 \leq \bar{\rho} \), \( \rho_1 = P(X_1 = 0) = \phi(0) \leq \bar{\rho} \), \( \rho_n \leq \bar{\rho}, \rho \leq \bar{\rho} \). \( \square \)
Corollary If $\mu := \sum_{k=0}^{\infty} kp_k \leq 1$ and $p_1 < 1$, the extinction happens with probability one. Indeed the derivative of the convex $\phi$ at 1 is smaller than 1. Or, for $x < 1$, if not all $p_2 = p_3 = \ldots = 0$,

$$1 > \sum_{k=1}^{\infty} \left( \sum_{i=0}^{k-1} x^i \right) p_k \geq \left( 1 - x \right)^{-1} \sum_{k=0}^{\infty} \left( 1 - x^k \right) p_k,$$

$$1 - x > \sum_{k=0}^{\infty} \left( 1 - x^k \right) p_k = 1 - \phi(x).$$

However, if $p_2 = p_3 = \ldots = 0$, then $\mu = p_1 < 1$, and the above computations are still valid.

Remark $\mu < 1$ implies that $p_1 < 1$.

If $\mu > 1$ there is positive probability to avoid extinction. Cases: (i) $p_0 = 0$, then $p(1) < 1$ and $X_n$ is increasing $X_n \to \infty$ as $n \to \infty$; (ii) $p(0) > 0$, but the chain is transient spends finite time in any finite set, hence, $X_n \to \infty$ as $n \to \infty$ with positive probability, albeit there is positive probability of extinction.

END OF MARKOV CHAINS.

Lecture 9

Ch 2. Poisson process

2.1. Exponential distribution.

A random variable $T$ is said to have an exponential distribution with rate $\lambda > 0$ if

$$P(T > t) = e^{-\lambda t}$$

for all $t \geq 0$. The density function

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & \text{for } t \geq 0, \\ 0 & \text{for } t < 0. \end{cases}$$

Compute $\mathbb{E}T = 1/\lambda$, $\mathbb{E}T^2 = 2/\lambda^2$, $\text{Var}T = 1/\lambda^2$.

Note

$$\int_0^{\infty} t^k e^{-t} dt = k! \lambda^{k+1}.$$

Observe $T\lambda = \text{expon}(1)$.

Why is it called rate?

$$P(T \in [s, s + t] \mid T > s) = \ldots \sim \lambda t.$$

Lack of memory property.

$$P(T > t + s \mid T > t) = P(T > s).$$
Electric bulbs. It is believed that the chances that an electric bulb burns off on a small time interval \([t, t + dt]\) are \(\lambda dt\) provided that it was working at time \(t\), that is
\[
P(t \leq T \leq t + dt \mid T \geq t) = \lambda dt,
\]
\[
P(T > t) - P(T > t + dt) = P(T > t)\lambda dt, \quad f' = -\lambda f, \quad P(T > t) = e^{-\lambda t}.
\]

An event happening with rate \(\lambda\)...

Exponential races. Let \(S = \text{expon}(\lambda)\), \(T = \text{expon}(\mu)\) and let them be independent.

\[
\min(S, T) = \text{expon}(\lambda + \mu).
\]

Joint density...

\[
E(T - S)_+ = \lambda \mu \int_0^\infty \int_s^\infty e^{-\lambda s}e^{-\mu(t-s)} ds dt = \frac{\lambda}{\mu} \int_0^\infty e^{-(\lambda+\mu)s} ds = \frac{\lambda}{\mu(\lambda + \mu)}.
\]

Exponential races of \(n\) random variables

**Theorem 2.1** Let \(T_i\) be independent exponential(\(\lambda_i\)) random variables given for \(i = 1, \ldots, n\). Define \(V = \min\{T_1, \ldots, T_n\}\) and let \(I\) be such that \(T_I = V\). Then
\[
P(V > t) = \exp(-(\lambda_1 + \cdots + \lambda_n)t),
\]
\[
P(I = i) = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n},
\]
and \(V\) and \(I\) are independent.

Proof. The first is very easy.

Then use the case \(n = 2\), take \(i\) set \(S = T_i\), \(T = \min_{j \neq i} T_j\). Then
\[
P(I = i) = P(S < T)...
\]

\[
P(I = i, V > v) = P(v < S < T) = \int_v^\infty \int_s^\infty \lambda_i \mu_i e^{-(\lambda_i s + \mu_i t)} ds dt = \frac{\lambda_i}{\lambda_i + \mu_i} e^{-(\lambda_i + \mu_i)v},
\]
hence independence. \(\square\)

**Example 2.2.** A submarine has three navigational devices but can remain at sea if at least two are working. Suppose that the failure times are exponentially distributed with means 1 year, 1.5 year, and 3 years. What is the average length of time the submarine can remain at sea? (Wrong answer in the book)

We have \(\lambda_1 = 1\), \(\lambda_2 = 2/3\), \(\lambda_3 = 1/3\), \(\lambda_1 + \lambda_2 + \lambda_3 = 2\). Let \(\tau_i\) be the failure time of the \(i\)th.
We need to find
\[ E(\tau_2 \land \tau_3)I_{\tau_1 < \tau_2 \land \tau_3} + E(\tau_1 \land \tau_3)I_{\tau_2 < \tau_1 \land \tau_3} + E(\tau_2 \land \tau_1)I_{\tau_3 < \tau_2 \land \tau_1}. \]

We have
\[
E(\tau_2 \land \tau_3)I_{\tau_1 < \tau_2 \land \tau_3} = E \tau_1 I_{\tau_1 < \tau_2 \land \tau_3} + E(\tau_2 \land \tau_3 - \tau_1)+
\]
\[
= EVI_{I=1} + \frac{1}{1+1} = (1/2)(1/2) + 1/2 = 3/4,
\]
\[
E(\tau_1 \land \tau_3)I_{\tau_2 < \tau_1 \land \tau_3} = E \tau_2 I_{\tau_2 < \tau_1 \land \tau_3} + E(\tau_1 \land \tau_3 - \tau_2)+
\]
\[
= EVP(I = 2) + \frac{2/3}{(4/3)(2/3 + 4/3)} = 1/6 + 1/4 = 5/12,
\]
\[
E(\tau_2 \land \tau_1)I_{\tau_3 < \tau_2 \land \tau_1} = E \tau_3 I_{\tau_3 < \tau_2 \land \tau_1} + E(\tau_2 \land \tau_1 - \tau_3)+
\]
\[
= EVP(I = 3) + \frac{1/3}{(5/3)(1/3 + 5/3)} = 1/12 + 1/10 = 11/60.
\]

The answer is (different from the textbook)
\[
(45 + 25 + 11)/60 = 81/60 \text{ (years)}.
\]

2.2. Defining the Poisson process.

Take \( \lambda > 0 \).

Definition We say that \( X \) has a Poisson distribution with parameter \( \lambda \), or \( X = \text{Poisson}(\lambda) \), if
\[
P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \quad n = 0, 1, 2, \ldots
\]

Theorem 2.3. p.100 For any \( k \geq 1 \)
\[
EX(X - 1)\ldots(X - k + 1) = \lambda^k.
\]

In particular, \( EX = \text{Var} \ X = \lambda \).

Proof.
\[
\phi_X(s) = E s^X = e^{-\lambda(1-s)}
\]
Lecture 10. Theorem 2.4. p.100 If \(X_i\) are independent Poison(\(\lambda_i\)), \(X = X_1 + \ldots + X_n\) is Poison(\(\lambda = \sum_i \lambda_i\)).

Proof.

\[ \phi_X(s) = e^{-\lambda(1-s)}. \]

Theorem. (Thining a Poisson random variable) Let \(N\) be \(\lambda\) Poisson and \(Y_i, i = 1, \ldots,\) be iid \(A = \{a_1, a_2, \ldots, a_m\}\)-valued random variables. Introduce \(N_j\) as the number of \(i \leq N\) such that \(Y_i = a_j, j = 1, 2, \ldots\) Then the \(N_j\) are independent \(\lambda P(Y_1 = a_j) = \lambda p_j\) Poisson.

Proof. Fix \(k_1, \ldots, k_m\) and introduce \(A_j = \{\#(i \in \{1, \ldots, m\} : Y_i = a_j) = k_j\}\).

We have

\[
P(N_1 = k_1, \ldots, N_m = k_m) = P(N = \sum k_i)P(A_1 \ldots A_m).
\]

\[
e^{-\lambda} \frac{\lambda^{k_1+\ldots+k_m}}{(k_1+\ldots+k_m)!} \frac{(k_1+\ldots+k_m)!}{k_1!\ldots k_m!} p_1^{k_1} \ldots p_m^{k_m} = \prod_{j=1}^m \frac{(\lambda p_j)^{k_j}}{k_j!} e^{-\lambda p_j}.
\]

Example 2.8. Suppose that in each day of the year there is a basketball game and the number of people to attend each of these games is 2263 Poison independent random variables. What is the probability that in each day in the year at least one of the people in attendance has that birthday?

Let \(Y_i\) be iid \(\{0, 1\}\)-valued random variables with \(P(Y_i = 1) = 1/365\). The number of people out of \(N\) whose birthday is on day one is \(2263/365 = 6.2\) Poisson. At least one has that day birthday is

\[1 - e^{-6.2}.\]

The answer is

\[(1 - e^{-6.2})^{365} = 0.4764.\]

Definition Let \(\tau_1, \tau_2, \ldots\) be iid expon(\(\lambda\)) random variables. Introduce \(T_n = \tau_1 + \ldots + \tau_n, T_0 = 0\). Then

\[N(s) = \max(n : T_n \leq s)\]

is called a rate \(\lambda\) Poisson process.

Light bulbs on real line (or students flipping a coin to decide to go to the bank): \(\lambda dt\) the probability for each sell to light up. How far from zero the first light is \((t = k dt)\)

\[P(\tau_1 > t) = (1 - \lambda dt)^{t/dt} = e^{-\lambda t}.
\]

Theorem* The increments \(N(t_1) - N(t_0), \ldots, N(t_n) - N(t_{n-1}), 0 \leq t_0 < \ldots < t_n\) are independent and, for any \(s \geq 0\), the process \(N(t + s) - N(s), t \geq 0\), is independent of \(N(r), r \in [0, s]\).
Proof. The number of shining bulbs on the intervals \((t_0, t_1], \ldots, (t_{n-1}, t_n]\) are independent.

**Lemma 2.7.** For any \(s \geq 0\), the process \(N(t + s) - N(s), t \geq 0\), is a rate \(\lambda\) Poisson process.


A paradox. Imagine the model with light bulbs on the whole real line. What is the average length of the interval with dark bulbs covering the origin. Interpretation with joining a line with exponential rate of service.

**Lemma 2.6.** p. 103 For \(t > s \geq 0\), \(N(t) - N(s)\) is Poisson\([\lambda(t - s)]\).

Proof. By Lemma 2.7 and Theorem 2.2 for \(\tau = t - s\) we have

\[
P(N(t) - N(s) < n) = P(N(t - s) < n) = P(T_n > t - s) = \int_{t-s}^{\infty} \frac{(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} \, dx
\]

\[
= \int_{0}^{\infty} \frac{\lambda y}{(n-1)!} e^{-\lambda y} \, dy
\]

\[
= e^{-\lambda \tau} \sum_{k=0}^{n-1} \frac{(\lambda \tau)^{n-k+1}}{k!(n - k + 1)!} \int_{0}^{\infty} \frac{(\lambda y)^k}{k!} e^{-\lambda y} \, d\lambda y = e^{-\lambda \tau} \sum_{k=0}^{n-1} \frac{(\lambda \tau)^{n-k+1}}{(n - k + 1)!}
\]

\[
= e^{-\lambda \tau} \sum_{k=0}^{n-1} \frac{(\lambda \tau)^k}{k!}.
\]

**Theorem.** If \(N(s)\) is a rate \(\lambda\) Poisson process, then

(i) \(N(0) = 0\), \(N(t)\) is right continuous,

(ii) \(N(t + s) - N(s)\) is Poisson\((\lambda t)\),

(iii) \(N(t)\) has independent increments.

Conversely, if (i), (ii), and (iii) hold, then \(N(t)\) is a rate \(\lambda\) Poisson process.

**Lecture 11**

Proof. Converse. First observe that for any \(t > 0, s_1, \ldots, s_n \leq t \leq t_1, \ldots, t_m\)
\((N(s_1), \ldots, N(s_n))\) is independent of \((N(t_1) - N(t), \ldots, N(t_m) - N(t))\). Then

**Claim** Define \(\tau = \inf\{t \geq 0 : N_t \geq 1\}\). Then

\[
P(\tau > t) = P(N(t) = 0) = e^{-\lambda t};
\]

the process \(N(t + \tau) - N(\tau)\) is independent of \(\tau\), possesses the properties (i)-(iii), and \(N_\tau = 1\).

To prove the claim define \(\tau^n = (k + 1)/n\) if \(k/n < \tau \leq (k + 1)/n\). The events

\[
\{\tau^n = (k + 1)/n\} = \{N(k/n) = 0, N((k + 1)/n) \geq 1\}
\]

and

\[
\{N(\tau^n + t_1) - N(\tau^n) = k_1, \ldots, N(\tau^n + t_m) - N(\tau^n) = k_m\}
\]

are independent, so \(\tau^n\) and the process \(N(\tau^n + t) - N(\tau^n)\) are independent.

To prove that \(N_\tau = 1\) observe that (!!!)

\[
P(N_\tau \geq 2, \tau \leq T)
\]
$$= \sum_{k=0}^{n-1} P(Tk/n < \tau \leq T(k+1)/n, N(Tk/n) = 0, N(T(k+1)/n) \geq 2)$$

$$\leq \sum_{k=0}^{n-1} P(N(T/n) \geq 2) = ne^{-\lambda T/n} \sum_{j=2}^{\infty} \frac{(\lambda T/n)^j}{j!} \to 0$$

Then induction on the number of jumps. \hfill \Box

How the Poisson process may arise in practice, \( n \) customers, each one may come with probability \( \lambda/n \) during the time interval \((0, 1)\) choosing time to come to be uniformly distributed on \((0, 1)\)

The probability of having exactly \( k \) people coming is

$$p_{nk} := \frac{n(n-1)...(n-k+1)}{n!} (\lambda/n)^k (1 - \lambda/n)^{n-k}.$$ 

**Theorem 2.5** \( p_{nk} \to e^{-\lambda} \lambda^k / k! \)

The probability of having exactly \( k \) people coming during a fixed interval of time \( \Delta \subset (0, 1) \) of length \( |\Delta| \) tends to \( e^{-\lambda|\Delta|} (\lambda|\Delta|)^k / k! \).

Given two disjoint \( \Delta_1, \Delta_2 \subset (0, 1) \), the probability of having exactly \( k_1 \) and \( k_2 \) people coming during \( \Delta_1 \) and \( \Delta_2 \) respectively, is given by the multinomial distribution

$$p_{nk_1k_2} = \frac{n!}{k_1!k_2!(n-k_1-k_2)!} \left( \frac{\lambda|\Delta_1|}{n} \right)^{k_1} \left( \frac{\lambda|\Delta_2|}{n} \right)^{k_2} \left( 1 - \frac{\lambda(|\Delta_1| + |\Delta_2|)}{n} \right)^{n-k_1-k_2}$$

which tends to

$$e^{-\lambda|\Delta_1|} (\lambda|\Delta_1|)^{k_1} / k_1! e^{-\lambda|\Delta_2|} (\lambda|\Delta_2|)^{k_2} / k_2!$$

meaning in particular asymptotical independence of the numbers of customers.

2.4.1. Thinning.

**Theorem 2.11** Let \( N(t) \) be a rate \( \lambda \) Poisson process and \( Y_i, i = 1, ..., \) be iid \( A = \{a_1, a_2, ..., a_m\} \)-valued random variables. Introduce \( N_j(t) \) as the number of \( i \leq N(t) \) such that \( Y_i = a_j \), \( j = 1, 2, ..., m \). Then the \( N_j(t) \) are independent \( \lambda P(Y_1 = a_j) \) Poisson processes.

Proof. What does the independence of processes mean? Given processes \( N_1(t), ..., N_m(t) \), they are independent if for any \( 0 \leq t_0 < t_1 < ... < t_n \) the vectors

$$(N_1(t_0), ..., N_1(t_n)), ..., (N_m(t_0), ..., N_m(t_n))$$

are independent. If \( N_i(0) = 0 \) equivalent definition is that the vectors

$$(N_1(t_1) - N_1(t_0), ..., N_1(t_n) - N_1(t_{n-1})), ..., (N_m(t_1) - N_1(t_0), ..., N_m(t_n) - N_m(t_{n-1}))$$

are independent.

In light of Theorem*, for any \( 0 \leq t_0 < t_1 ... < t_n \), the vectors

$$(N_1(t_{i+1}) - N_1(t_i), ..., N_m(t_{i+1}) - N_m(t_i)), \ \ i = 0, 2, ..., n - 1,$$

are independent. Therefore, it suffices to check that the coordinates of these vectors are independent.
But these are obtained by thinning the rate \( \lambda(t_{i+1} - t_i) \) Poisson random variable \( N(t_{i+1}) - N(t_i) \) so that the result follows from Theorem on page 22.

More details: Fix \( k_1, ..., k_m \) and introduce

\[
A_j = \{ \#(i \in \{1, ..., m\} : Y_i = a_j) = k_j \}.
\]

We have

\[
P((N_1(t) - N_1(s), ..., N_m(t) - N_m(s)) = (k_1, ..., k_m)) \\
= P(N(t) - N(s) = k_1 + ... + k_m)P(A_1...A_m) \\
= e^{-\lambda(t-s)}(\lambda(t-s))^{k_1+...+k_m} \frac{(k_1 + ... + k_m)!}{k_1!...k_m!}p_1^{k_1}...p_m^{k_m} = \prod_{j=1}^{m} \frac{(\lambda p_j)^{k_j}}{k_j!} e^{-\lambda p_j(t-s)}.
\]

\( \square \)

Splitting one Poisson process by using an iid sequence is called thinning. Going in the other direction is called superposition.

Nonhomogeneous Poisson Processes. We say that \( N(s) \) is a Poisson Process with rate \( \lambda(r) \) if

(i) \( N(0) = 0 \) and \( N(t) \) is right continuous,

(ii) \( N(t) \) has independent increments,

(iii) \( N(t) - N(s) \) is Poisson with mean \( \int_s^t \lambda(r) \, dr \).

Distribution of the first jump of \( N ... \)

Denoting by \( \phi(t) \) the inverse function to \( \Lambda(t) := \int_0^t \lambda(r) \, dr \) so that

\[
\phi_t = \int_0^t \lambda^{-1}(\phi(r)) \, dr
\]

and letting \( M(t) = N(\phi(t)) \) we have

(i) \( M(0) = 0 \) and \( M(t) \) is right continuous,

(ii) \( M(t) \) has independent increments,

(iii) \( M(t) - M(s) \) is Poisson with mean \( t - s \).

If \( T_i \) are the jump times of \( M \) then \( \Lambda(T_i) \) are the jump times of \( N \). \( T_{i+1} - T_i \) are exponential independent but not so \( \Lambda(T_{i+1}) - \Lambda(T_i) \).

Example 2.9 \( (M/G/\infty \text{ Queue}) \) Every student is talking on their smartphone. The arrivals at the ATM follow a Poisson process. As for the duration of the calls themselves, there is no reason to suppose that it has an exponential distribution, so we use a general distribution function \( G \) with \( G(0) = 0 \) and mean \( \mu \). We are interested in the number of busy lines at time \( t \), assuming that there were no busy lines at time 0.

Theorem 2.12. Time nonhomogeneous case. Suppose that in a Poisson process with rate \( \lambda \) we keep a point that lands at \( s \) with probability \( p(s) \). Then the result is a nonhomogeneous Poisson process with rate \( \lambda p(s) \).
Doing Example 2.9. For any $t$, by Theorem 2.12 the number of calls still in progress is Poisson with mean
\[
\int_0^t \lambda (1 - G(t - s)) \, ds = \lambda \int_0^t (1 - G(s)) \, ds.
\]
As $t \to \infty$ Theorem 2.1.3 In the long run the number of calls in the system will be Poisson with mean
\[
\lambda \int_0^\infty (1 - G(s)) \, ds = \lambda \mu.
\]

Lecture 12

2.3. Compound Poisson process, p. 106.

Example 2.3. Consider the McDonald’s restorant on Route 13 in Ithaca. Assume that between 12:00 and 13:00 cars arrive according to a Poisson process with rate $\lambda$. Let $Y_i$ be the number of people in the car number $i$. Assume the $Y_i$ are iid and independent of the arrival times.

Let $S(t) = Y_1 + \ldots + Y_{N(t)}$ the number of customers that arrived up to time $t$. Mean, variance of $S(t)$?

Theorem 2.10. Let $Y_i$ be iid, let $N$ be an independent nonnegative integer-valued random variable, and let $S = Y_1 + \ldots + Y_N$ with $S = 0$ if $N = 0$ Then

(i) If $E|Y_i|, \, EN < \infty$, then $ES = EN EY_i$;
(ii) If $E|Y_i|^2, \, EN^2 < \infty$, then $\text{Var} S = EN \text{Var}(Y_i) + \text{Var}(N)(EY_i)^2$;
(iii) If $N$ is Poisson $\lambda$, then $\text{Var} S = \lambda EY_i^2$.

Proof. We have
\[
Ef(S, N) = \sum_{n=0}^\infty Ef(Y_1 + \ldots + Y_n, n) P(N = n) = \sum_{n=0}^\infty Ef(Y_1 + \ldots + Y_n, n) P(N = n).
\]

If $EY_i = 0$
\[
ES^2 = EY_i^2 \sum_{n=0}^\infty n P(N = n) = EN EY_i^2.
\]

In the general case $Y_i = Z_i + m$, where $m = EY_i$ and
\[
ES^2 = E(Z_1 + \ldots + Z_N + mN)^2 = E(Z_1 + \ldots + Z_N)^2 + m^2 EN^2
\]
\[
= EN \text{Var}(Y_i) + m^2 EN^2.
\]

Example 2.5. p. 108. Suppose that the number of customers at a liquor store has a Poisson distribution with mean 81 and that each customer spends an average of $8 with a standard deviation of $6. It follows from the theorem that the mean revenue for the day is 81$8=$648. The variance of that revenue is
\[
81((\$6)^2 + (\$8)^2) = 8100 \text{ (square dollars)}
\]
and the standard deviation is $90 with a mean revenue of $648.
Theorem 2.14. p. 112. Let $N_1(t), ..., N_k(t)$ be independent Poisson processes with rates $\lambda_1, ..., \lambda_k$. Then $N(t) = N_1(t) + ... + N_k(t)$ is Poisson with rate $\lambda := \lambda_1 + ... + \lambda_k$.

Proof. Take one with rate $\lambda$ and split... $p_i = \lambda_i/\lambda$. \square

Example 2.12 (A Poisson race.) Red arrivals with rate $\lambda$ and independent green arrivals with rate $\mu$. What is the probability to have 6 red arrivals before a total of 4 green arrivals?

Out of the first 9 arrivals at least 6 are red. Answer

$$\sum_{k=6}^{9} \binom{9}{k} p^k (1-p)^{9-k}, \quad p = \frac{\lambda}{\lambda + \mu}.$$ 

2.4.3. Conditioning. p. 113.

Let $T_1, T_2, ...$ be the arrival times of a rate $\lambda$ Poisson process, let $U_1, ..., U_n$ be iid uniformly distributed on $[0, t]$, and let $V_1 < V_2 < ... < V_n$ be the $U_i$, rearranged into increasing order.

Teorem 2.15. The conditional distribution of $(T_1, T_2, ..., T_n)$ given $N(t) = n$ coincides with the distribution of $(V_1, V_2, ..., V_n)$.

Proof. Light bulbs. Given that on $[0, t]$ exactly $n$ bulbs burning, if they light up with equal probability. The positions of the first, second,... the light ups have the same uniform distribution. \square

Corollary Teorem 2.16.If $s < t$ and $0 \leq m \leq n$. then

$$p(n.m, s, t) := P(N(s) = m \mid N(t) = n) = \binom{n}{m} \left( \frac{s}{t} \right)^m \left( 1 - \frac{s}{t} \right)^{n-m}.$$ 

Different derivation. We have

$$p(n.m, s, t) = \frac{P(N(t) = n, N(s) = m)}{P(N(t) = n)} = \frac{P(N(t) - N(s) = n - m \mid N(s) = n)P(N(s) = n)}{P(N(t) = n)}.$$ 

Example 2.14. p.115. Trucks and cars on highway US 421 are Poisson processes with rate 40 and 100 per hour, respectively 1/8 of the trucks and 1/10 of the cars get off on exit 257 to go to town A. Find

(a) Find the probability that exactly six trucks arrive at A between noon and 1 p.m.
(b) Given that there were six truck arrivals at A between noon and 1 p.m., what is the probability that exactly two arrived between 12:20 and 12:40?
(c) If we start watching at noon, what is the probability that at least four cars arrive before two trucks do?
(d) Suppose that all trucks have 1 passenger while 30% of cars have 1 passenger, 50% have 2, and 20% have 4. Find the mean and the standard deviation of the number of customers that arrive at A in one hour.

Solution.
(a) Trucks get off with rate $\mu = 40/8 = 5$. The answer is $e^{-5}5^6/6! = 0.1462$.
(b) By conditioning: $\binom{6}{2}(1/3)^2(2/3)^4 = 0.3292$.
(c) By Poisson race: $\lambda = 10, \mu = 5$.

$$P(\text{when 5 vehicles arrive, all are cars}) = \left(\frac{\lambda}{\lambda + \mu}\right)^5,$$

$$P(\text{when 5 vehicles arrive, 4 are cars and 1 is a truck}) = 5\left(\frac{\lambda}{\lambda + \mu}\right)^4 \left(\frac{\mu}{\lambda + \mu}\right).$$

(d) We have $N(1)$ Poisson $\lambda + \mu, Y_1 + \ldots + Y_{N(1)}$,

$$P(Y_i = 1) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} 0.3, \quad P(Y_i = 2) = \frac{\lambda}{\lambda + \mu} 0.5$$

$$P(Y_i = 3) = 0, \quad P(Y_i = 4) = \frac{\lambda}{\lambda + \mu} 0.2.$$
Then $S_n/n$ is again Cauchy with the same distribution.

Lecture 13
Proof of Theorem 3.1. By Theorem 3.2 we have

$$ \frac{T_n}{n} \to \mu. $$

Hence

$$ \frac{T_{N(t)+1}}{N(t)}, \frac{T_N(t)}{N(t)} \to \mu $$
as $t \to \infty$. Using $T_{N(t)} \leq t \leq T_{N(t)+1}$ we are done. \hfill \Box

Suppose when the $i$th renewal occurs we get a reward $r_i$, Assume $(t_i, r_i)$ are iid. Let

$$ R(t) = \sum_{i=1}^{N(t)} r_i. $$

**Theorem 3.3.** With probability one

$$ \frac{R(t)}{t} \to \frac{E r_i}{E t_i}. $$

Proof. We have

$$ \frac{R(t)}{t} = \left( \frac{1}{N(t)} \sum_{i=1}^{N(t)} r_i \right) \frac{N(t)}{t}. $$

The result is natural: the limiting rate of reward is the reward per cycle.

**Example 3.3.** Long run car costs. Lifetime of a car is a random variable $\tau$ with density $h$. We buy a new one once the old one breaks down or reaches $T$ years. The cost of a new car is $A$ dollars and an additional $B$ dollars is required to repair the vehicle if it breaks down before time $T$. What is the long-run cost per unit time of this policy?

Solution. The duration of the $i$th cycle, $t_i = \tau \wedge T$,

$$ Et_i = Et_i I_{t_i \leq T} + T P(t_i > T) = \int_0^T th(t) \, dt + T \int_T^\infty h(t) \, dt. $$

The cost of the $i$th cycle

$$ Er_i = A + B \int_0^T h(t) \, dt. $$

Concrete example $h$ is uniform on $[0, 10]$, $A = 10$ (thousand dollars), $B = 3$,

$$ \frac{1}{10} \int_0^T t \, dt = \frac{T^2}{20} \left( +T \frac{1}{10} (10 - T) \right), $$

$$ \frac{Er_i}{Et_i} = \frac{10 + 0.3T}{T - 0.05T^2}. $$
To minimize (not maximize) take the derivative

\[ T = \frac{-1 \pm \sqrt{1 + 4(0.015)(10)}}{2(0.015)} = \frac{-1 \pm \sqrt{1.6}}{0.03}. \]

\( T = 8.83 \) years.

Example 3.4. Alternating renewal processes. Suppose \( s_1, s_2, \ldots \) are iid \( \geq 0 \) with distribution \( F \) and mean \( \mu_F \) and let \( u_1, u_2, \ldots \geq 0 \) be iid with distribution \( G \) and mean \( \mu_G \) independent of \( s_i \). Our alternating process spends time \( s_i \) in state 1 and then spends time \( u_i \) in state 2 and then repeats the cycle again.

**Theorem 3.4.** The limiting fraction of time spent in state 1 is

\[ \frac{\mu_F}{\mu_F + \mu_G}. \]

Proof. Set \( t_i = s_i + u_i \) and consider the time spent at state 1 as a reward. Then Theorem 3.3 applies.

Example 3.5. Poisson janitor. A light bulb burns for amount of time having distribution \( F \) with mean \( \mu_F \). A janitor inspects the bulb at times forming a rate \( \lambda \) Poisson process and replaces the bulb if it is burnt out.

(a) At what rate are bulbs replaced? \( \lim(N(t)/t) =? \) \( N(t) \) is the number of bulbs replaced.

(b) What is the limiting fraction of time that the light bulbs work?

(c) What is the limiting fraction of visits on which the bulb is working?

Solution

Renewal process, \( \tau_i \)'s stand for interarrival times of the janitor, \( s_i \)'s stand for the life time of bulbs.

\[ t_1 = \tau_1 + \ldots + \tau_{n_1}, \quad n_1 = \min \{ n : \tau_1 + \ldots + \tau_n \geq s_1 \}. \]

(a) Small lights interpretation: \( t_1 = s_1 + u_1 \), where \( u_1 \) is \( \text{expon}(\lambda) \). Number of bulbs replaced per cycle is 1. The average length of the cycle \( Et_1 = \mu_F + 1/\lambda \). The rate \( 1/\text{Et}_1 \),

\[ \frac{N(t)}{t} \to \frac{1}{\mu_F + 1/\lambda}. \]

(b) Rewards \( s_i \)

\[ \frac{E{s_i}}{Et_i} = \frac{\mu_F}{\mu_F + 1/\lambda} \]

(c) Let \( V(t) \) be the number of visits until time \( t \). \( V(t)/t \to \lambda \). The limiting fraction of the number of inspections on which a bulb was replaced (not the question which was asked)

\[ \frac{N(t)}{V(t)} \to \frac{1}{\lambda(\mu_F + 1/\lambda)} \quad (1 - \frac{1}{\lambda\mu_F + 1} = \frac{\lambda\mu_F}{\lambda\mu_F + 1}). \]

Lecture 14
3.2. Application to queueing theory. p. 130.
3.2.1. GI/G/1 queue.
Times between successive arrivals $t_i$ have a distribution $F$ with mean $1/\lambda$. The long-run arrival rate

$$\lim_{t \to \infty} \frac{N(t)}{t} = \lambda.$$  

The $i$th customer requires service time $s_i$ iid with a distribution $G$ and mean $1/\mu$.

**Theorem 3.5.** Suppose $\lambda < \mu$. If the queue starts with some number $k \geq 1$ customers who need service, then it will empty out with probability one. Furthermore, the limiting fraction of time the server is busy is $= \lambda/\mu$.

**Proof.** Let $T_n = t_1 + ... + t_n$ be the time of the $n$th arrival. Then

$$\frac{T_n}{n} \to 1/\lambda.$$  

Let $Z_0$ be the sum of the service times of the customers in the system at time 0 and les $s_i$ be the service time of the $i$th customer to arrive after time 0. Set $S_n = s_1 + ... + s_n$. Then the time $Z_n$ to serve all customers in the system up to the $n$th customer is $Z_0 + S_n$ and

$$\frac{Z_0 + S_n}{n} \to 1/\mu.$$  

Since $1/\mu < 1/\lambda$, starting from an $n_0$, for $n \geq n_0$,

$$Z_0 + S_{n-1} < T_n$$  

and at the $n_0$th arrival there is no queue and the system starts out empty.

Assume it starts out empty in the beginning. Let $B_n$ be the amount of time the server has been busy up to time $T_n$

$$B_n = S_n - Z_n,$$  

where $Z_n$ is the amount of time to empty the system after time $T_n$ from customers in the system at time $T_n$. We have

$$\frac{B_n}{T_n} = \frac{S_n}{T_n} - \frac{Z_n}{T_n} \to \frac{\lambda}{\mu}.$$  

Note that $Z_n$ vanishes once in a while. Imaging the limit exists. \qed

### 3.2.2. Cost equations.

Let $X(s)$ be the number of customers in the system at time $s$. Let $W_m$ be the time the $m$th customer spends in the system, $N(t)$ be the number of customers who joined the system up to time $t$.

Let $L$ be the long-run time average number of customers in the system

$$L = \lim_{t \to \infty} \frac{1}{t} \int_0^t X(s) \, ds.$$  

Let $W$ be the long-run average amount of time a customer spends in the system

$$W = \lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^n W_m.$$
Let \( \lambda_a \) be the long-run average rate at which customers join the system

\[
\lambda_a = \lim_{t \to \infty} \frac{N(t)}{t}.
\]

**Theorem 3.6. Little’s formula.** \( L = \lambda_a W \).

**Explanation.** Suppose each customer pays $1 for each minute spent in the system. When \( l \) customers are in the system we are earning $W \), so in the long run we earn an average $L per minute. On the other hand, imagine the customers pay $W \) for their entire time to spend in the system when they arrive. Then we earn at rate \( \lambda_a W \) per minute, i.e. the rate at which customers enter the system multiplied by the average amount they pay.

**Example 3.6. Waiting time in the queue.** GI/G/1 queue. We are interested (not quite!) in the customer’s average waiting time in the queue, \( W_Q \). If \( W \) is the average time the customer spends in the system, then

\[
W_Q = W - E s_i.
\]

Let \( L_Q \) be the average queue length in equilibrium (the one being served does not count if there is one). Suppose the customer pays $1 in the queue and repeat the derivation of Little’s formula, then

\[
L_Q = \lambda_a W_Q.
\]

Also for \( L \) which is the average number of customers in the system we have

\[
L = \lambda_a W.\]

\( L_Q \) is 1 less than the number of customers in the system, except when there are no customers. So if \( \pi(0) \) is the probability of no customers

\[
L_Q = L - 1 + \pi(0).
\]

Hence

\[
\pi(0) = L_Q - (L - 1) = 1 - \lambda_a E s_i, \quad E s_i = 1/\mu, \quad \pi(0) = 1 - \frac{\lambda}{\mu}.
\]

Observe that this agrees with Theorem 3.5.

3.2.3. M/G/1 queue. p. 133. M=Markovian, the input is a rate \( \lambda \) Poisson process. \( X_n \) is the number of customers in the queue when the \( n \)th customer enters service process. Then

\[
a_k = \int_0^\infty \frac{(\lambda t)^k}{k!} e^{-\lambda t} dG(t), \quad k = 0, 1, ...
\]

is the probability that \( k \) customers arrive during a service period (the server is busy).

\[
\sum_k k a_k = \lambda E s_i = \lambda/\mu.
\]

Let \( \zeta_1, \zeta_2, ... \) be iid with \( P(\zeta_i = k) = a_k \). We think of \( \zeta_i \) as the number of customers to arrive during the \( i \)th service period. If \( X_n > 0 \), then

\[
X_{n+1} = X_n + \zeta_n - 1.
\]
Generally

\[ X_{n+1} = (X_n + \zeta_n - 1)^+ \]

Theorem 3.7. p.134. If \( \lambda < \mu \), then \( X_n \) is positive recurrent and \( E_0T_0 = 1/\pi(0) = \mu/(\mu - \lambda) \).

If \( \lambda = \mu \), then \( X_n \) is null recurrent.

If \( \lambda > \mu \), then \( X_n \) is transient.

Proof. Note, for \( \tau_0 \) defined as the first time \( X_n \) reaches 0, obviously,

\[ m(i) = E_i\tau_0 = Ci \quad (*) \]

for some constant \( C \) and all \( i \geq 0 \). The chain \( Y_n := X_{n\wedge \tau_0} \) is a Markov chain with transition probability function given by

\[ p(i, i + k) = p(\zeta = k + 1), \quad k \geq -1, i \geq 1, \quad p(0, 0) = 1. \]

Therefore, by the Markov property, for \( i \geq 1 \), \( m(i) \) satisfies

\[ Ci = m(i) = 1 + \sum_{k=-1}^{\infty} p(\zeta = k + 1)m(i + k) = 1 + \sum_{k=0}^{\infty} p(\zeta = k)C(i + k - 1) \]

\[ = 1 + C(i - 1) + CE\zeta. \quad (***) \]

Observe that, if \( Y_n \) starts from \( i = 1 \), it will reach 0 with probability one iff \( X_n \) is recurrent, and the time to reach 0 for \( Y_n \) has finite expectation iff \( X_n \) is positive recurrent.

If \( E\zeta = \lambda/\mu < 1 \), the solution of (***) is \( C = (1 - E\zeta)^{-1} \). If \( E\zeta \geq 1 \), the only solution \( C = \infty \) and the chain is not positive recurrent.

Assume \( E\zeta > 1 \), and let us show that \( X_n \) is transient (or that \( Y_n \) reaches 0 with probability strictly less than one). Introduce

\[ \phi(s) = \phi(\zeta) = \sum_{k=0}^{\infty} s^k a_k \]

and let \( s_0 \) be the root of \( s = \phi(s) \) that is strictly less than 1. Such a root exists since \( \phi(0) > 0 \). Then observe that for \( i \geq 1 \)

\[ E_is_0^{X_1} = \sum_{k=-1}^{\infty} s_0^{i+k}P(\zeta = k + 1) = s_0^{i-1}\phi(s_0) = s_0^i. \]

Let \( \tau_N \) be the first time \( Y_n \geq N \). We know that \( h(i) = P_i(\tau_0 < \tau_N) \) satisfies for \( i \in [1, N - 1] \)

\[ E_i h(X_1) = h(i). \]

Also \( h(i) \leq s_0^i \) for \( i = 0 \) and \( i \geq N \) By Lemma** on page 16, we have \( h(i) \leq s_0^i \) for all \( i \). In particular, \( h(1) < s_0 \). Letting \( N \to \infty \) we get

\[ P_1(\tau_0 < \infty) \leq s_0 < 1, \]

hence transiency.