# ISyE 6761 (Fall 2016)

## Stochastic Processes I

Prof. H. Ayhan
Georgia Institute of Technology

LaTeXer: W. Kong
http://wwkong.github.io
Last Revision: May 25, 2017

## Table of Contents



These notes are currently a work in progress, and as such may be incomplete or contain errors.

# ACKNOWLEDGMENTS:

Special thanks to *Michael Baker* and his LaTeX formatted notes. They were the inspiration for the structure of these notes.

**Abstract**

The purpose of these notes is to provide the reader with a secondary reference to the material covered in ISyE 6761.

**Errata**

Test 1 - October 13th

Test 2 - November 17th

Breakdown of Grading: Test 1 (30%), Test 2 (30%), Assignments (10%), Final (30%)

All material will be posted on t-Square

Instructor Office Hours: 329 Groseclose (T,W,Th) 12:30pm-1:30pm

TA (Rui Gao): 425E Main Building (F) 12:30pm-2:30pm

# 1 Probability Theory

**Definition 1.1.** A **stochastic process** is a collection of random variables $\{X(t) : t \in T\}$ defined on a common probability space indexed by $T$. For example $X(k)$ can be the number of customers in a service system at time $k$ or the number of arrivals to a queuing system during the $n^{th}$ interarrival time.

**Example 1.1.** (Non-negative integer valued random variables) Let $X$ be a random variable taking values $\{0, 1, 2, ..., \infty\}$. Define $p_k = P(X = k)$ for $k = 0, 1, 2, ...$ and $P(X < \infty) = \sum_{k=0}^{\infty} p_k$, $P(X = \infty) = p_\infty = 1 - \sum_{k=0}^{\infty} p_k$. Define

$$E(X) = \begin{cases} \infty & P(X = \infty) > 0 \\ \sum_{k=0}^{\infty} kp_k & P(X = \infty) = 0 \end{cases}$$

If $f : [0, 1, ..., \infty] \to [0, \infty]$. We can also define

$$E[f(x)] = \sum_{0 \le k \le \infty} f(k)p_k$$

If $f : [0, 1, ..., \infty] \to [-\infty, \infty]$. We can define

$$E[f^+(x)] = \sum_{0 \le k \le \infty} f^+(k)p_k, f^+ = \max[f, 0]$$
$$E[f^-(x)] = \sum_{0 \le k \le \infty} f^-(k)p_k, f^- = -\min[f, 0]$$
$$E[f(x)] = E[f^+(x)] - E[f^-(x)]$$

The expected value is finite if and only if $E[|f(x)|] < \infty$. We call the special transformation below **variance**:

$$Var(X) = E\left[(X - E(X))^2\right]$$

**Example 1.2.** (Binomial Random Variable) Denoted as $b(k; n, p)$, we have

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

with expectation:

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\ &= np \underbrace{\sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-k-1)!} p^k (1-p)^{n-k-1}}_{=1} \\ &= np \end{aligned}$$

and variance:

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ E(X^2) &= ... = n(n-1)p^2 + np \end{aligned}$$

and reducing gives us $Var(X) = np(1-p)$.

**Example 1.3.** (Poisson random variable) Denoted as $p(k;\lambda)$, we have

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k = 0, 1, 2, ...$$

$$E(X) = \lambda, Var(X) = \lambda$$

**Example 1.4.** (Geometric random variable) Denoted as $g(k;p)$ and counting as the number of failures before the first success, we have

$$P(X = k) = (1-p)^k p, k = 0, 1, 2, ...$$

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k(1-p)^k p = \frac{1-p}{p} \\ Var(X) &= \frac{1-p}{p^2} \end{aligned}$$

**Lemma 1.1.** *If $X$is an integer valued non-negative random variable then $E(X) = \sum_{k=0}^{\infty} P(X > k)$.*

*Proof.* By direct evaluation:

$$\sum_{k=0}^{\infty} P(X > k) = \sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} P(X = j) = \sum_{j=1}^{\infty} P(X = j) \sum_{k=0}^{j-1} 1 = \sum_{j=1}^{\infty} jP(X = j)$$

□

In the multivariate case we have a random vector with non-negative integer valued components $\boldsymbol{X} = (X_1, ..., X_n)$ with joint distribution

$$P(X_1 = k_1, ..., X_n = k_n) = p_{k_1,...,k_n}$$

If $f$ attains non-negative values, then

$$E(f(\boldsymbol{X})) = \sum_{(k_1,...,k_n)} f(k_1, ..., k_n) p_{k_1,...,k_n}$$

If $f$ attains values in the real line, then

$$E[f(\boldsymbol{X})] = E[f^+(\boldsymbol{X})] - E[f^-(\boldsymbol{X})]$$

*Remark* 1.1. (Properties of the expected value and variance)

1) For $a_1, ..., a_n \in \mathbb{R}$, $E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i E[X_i]$

2) If $X_1, ..., X_n$ are independent random variables and $f_1, ..., f_n$ are bounded functions, then $E\left[\prod_{i=1}^n f_i(X_i)\right] = \prod_{i=1}^n E\left[f_i(X_i)\right]$

3) If $E[X_i^2] < \infty$ for $i = 1, ..., n$ and $Cov(X_i, X_j) = 0$ for all $i = 1, ..., n$ and $j = 1, ..., n$ then $Var(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^n Var(X_i)$

## 1.1 Convolution

Suppose $X$ and $Y$ are independent non-negative integer valued random variables with $P(X = k) = a_k$ and $P(Y = k) = b_k$. Then,

$$\begin{aligned} P(X+Y=n) &= \sum_{k=0}^n P(X=k, Y=n-k) \\ &= \sum_{k=0}^n a_k b_{n-k} \end{aligned}$$

**Definition 1.2.** The **convolution** of two sequences $\{a_n\}$ and $\{b_n\}$ is the new sequence $\{c_n\}$ where the $n^{th}$ element $c_n$ is defined by

$$c_n = \sum_{k=0}^n a_k b_{n-k}$$

We write $\{c_n\} = \{a_n\} * \{b_n\}$. Denote $\{p_k\} * ... * \{p_k\} = \{p_k\}^{n*} = p_k^{n*}$.

**Example 1.5.** Suppose $X$ is a $p(k;\lambda)$ random variable and $Y$ is a $p(k;\mu)$ random variable. Suppose $X$ and $Y$ are independent. Then, $X+Y$ is a $p(k;\lambda+\mu)$ random variable. The proof is as follows:

$$\begin{aligned} P(X+Y=n) &= \sum_{k=0}^n P(X=k)P(Y=n-k) \\ &= \sum_{k=0}^n \frac{e^{-\lambda}\lambda^k}{k!} \frac{e^{-\mu}\lambda^{n-k}}{(n-k)!} \\ &= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!} \sum_{k=0}^n \binom{n}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k} \\ &= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^n}{n!} \end{aligned}$$

**Example 1.6.** If $X$ is a $b(k;n,p)$ and $Y$ is a $b(k;m,p)$ and $X$ and $Y$ are independent. Then $X+Y$ is $b(k;n+m,p)$.

*Remark* 1.2. (Some properties of convolution)

1) Convolution of two probability mass functions is a probability mass function.

2) $X + Y \overset{d}{=} Y + X$ (equal in distribution; commutative)

3) $X + (Y + Z) \overset{d}{=} (X + Y) + Z$ (associative)

## 1.2 Generating Functions

**Definition 1.3.** Let $a_0, a_1, a_2...$ be a numerical sequence. If there exists $s_0 > 0$ such that $A(s) = \sum_{k=0}^\infty a_k s^k$ converges in $|s| < s_0$, then we call $A(s)$ the **generating function** of the sequence $\{a_n\}$. If $\{p_k : k \geq 0\}$ then $P(s) = \sum_{k=0}^\infty p_k s^k = E[s^X]$. If $\sum_{k=0}^\infty p_k = 1$ then $P(1) = 1$.

**Example 1.7.** If $X$ is $p(k;\lambda)$ then

$$P(s) = \sum_{k=0}^\infty \frac{e^{-\lambda}\lambda^k}{k!} s^k = e^{\lambda(s-1)}, \forall s > 0$$

If $X$ is $b(k;n,p)$,

$$P(s) = \sum_{k=0}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} s^k = (1-p+ps)^n$$

If $X$ is $g(k;p)$,

$$P(s) = \sum_{k=0}^{\infty} (1-p)^k p s^k = \frac{p}{1-(1-p)s}, \forall s < \frac{1}{1-p}$$

*Remark* 1.3. Note that

$$\frac{d^n}{ds^n} P(s) = \sum_{k=n}^{\infty} k(k-1)...(k-n+1) p_k s^{k-n} = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} p_k s^{k-n}$$

and

$$\frac{d^n}{ds^n} P(s)\Big|_{s=0} = n! p_n$$

**Proposition 1.1.** *The probability generating function uniquely defines its probability mass function.*

**Proposition 1.2.** *Let $X$ have a probability mass function with $p_k = P(X = k)$ and $\sum_{k=0}^{\infty} p_k = 1$. Let $q_k = P(X > k)$ and define $Q(s) = \sum_{k=0}^{\infty} q_k s^k$. Then*

$$Q(s) = \frac{1-P(s)}{1-s}, \forall s \in (0,1)$$

*Proof.* By direct evaluation,

$$\begin{aligned} Q(s) &= \sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} p_j s^k = \sum_{j=1}^{\infty} p_j \sum_{k=0}^{j-1} s^k \\ &= \sum_{j=1}^{\infty} p_j \frac{1-s^j}{1-s} = \frac{1}{1-s} \left( \sum_{j=1}^{\infty} p_j - \sum_{j=1}^{\infty} p_j s^j \right) \\ &= \frac{1}{1-s}(1 - p_0 - P(s) + p_0) \\ &= \frac{1-P(s)}{1-s} \end{aligned}$$

□

*Remark* 1.4. By the Monotone Convergence Theorem,

$$\lim_{s\to 1} Q(s) = \lim_{s\to 1} \sum_{k=0}^{\infty} q_k s^k = \sum_{k=0}^{\infty} \lim_{s\to 1} q_k s^k = \sum_{k=0}^{\infty} q_k = E[X]$$

*Remark* 1.5. By direct evaluation,

$$\begin{aligned} \frac{d}{ds} P(s)\Big|_{s=1} &= \sum_{k=1}^{\infty} k p_k = E[X] \\ \frac{d^2}{ds^2} P(s)\Big|_{s=1} &= E[X(X-1)] \\ &\vdots \\ \frac{d^n}{ds^n} P(s)\Big|_{s=1} &= E[X(X-1)...(X-n+1)] \end{aligned}$$

**Example 1.8.** If $X$ is $g(k;p)$ then

$$P(s) = \frac{p}{1-(1-p)s} \implies \frac{d}{ds} P(s) = \frac{p(1-p)}{(1-(1-p)s)^2} \implies \frac{d}{ds} P(s)\Big|_{s=1} = \frac{p(1-p)}{p^2} = \frac{1-p}{p}$$

*Remark* 1.6. Note that $Var(X) = P''(1) + P'(1) - (P'(1))^2$.

*Remark* 1.7. The generating function of the sum of independent random variables is the product of their generating functions.

(1) Formally, if $X_i$ for $i = 1, 2$ are independent non-negative integer valued random variables with generating functions

$$P_{X_i}(s) = E\left[s^{X_i}\right], i = 1, 2$$

and $0 \leq s \leq 1$ then

$$P_{X_1+X_2}(s) = E\left[s^{X_1+X_2}\right] = E[s^{X_1}]E[s^{X_2}] = P_{X_1}(s)P_{X_2}(s)$$

(2) If $\{a_j\}$ and $\{b_j\}$ are two sequences with generating functions $A(s), B(s)$ then the generating functions of $\{a_n\} * \{b_n\}$ is $A(s)B(s)$. This is obvious from the definition:

$$\begin{aligned}
\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) s^n &= \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} a_k b_{n-k} s^n \\
&= \sum_{k=0}^{\infty} a_k s^k \sum_{n=k}^{\infty} b_{n-k} s^{n-k} \\
&= A(s)B(s)
\end{aligned}$$

**Example 1.9.** If $X_1, X_2$ are respectively $p(k;\lambda), p(k;\mu)$ and $X_1$ and $X_2$ are independent, then

$$P_{X_1+X_2}(s) = e^{(\lambda+\mu)(s-1)}$$

which is the generating function of a $p(k;\lambda+\mu)$.

**Example 1.10.** Suppose that $X_1, ..., X_n$ are independent and identically distributed (iid) random variables with

$$X_i = \begin{cases} 1 & \text{with } p \\ 0 & \text{with } (1-p) \end{cases}, i = 1, ..., n$$

Then

$$P_{X_i}(s) = ps + (1-p), P_{X_1+...+X_n}(s) = (ps + (1-p))^n$$

*Remark* 1.8. (Random sums of random variables) Consider iid non-negative random variables $\{X_n : n \geq 1\}$ with $p_k = P(X_1 = k), P_{X_1}(s) = E[s^{X_1}]$. Let $N$ be independent of $\{X_n : n \geq 1\}$ and suppose that $P(N = j) = \alpha_j$ for $j = 0, 1, 2, ....$ Define

$$s_0 = 0, s_1 = X_1, ...., s_N = X_1 + ... + X_N$$

From conditional probability,

$$\begin{aligned}
P(S_n = j) &= \sum_{k=0}^{\infty} P(S_N = j | N = k)P(N = k) \\
&= \sum_{k=0}^{\infty} P(S_k = j)P(N = k) \\
&= \sum_{k=0}^{\infty} p_j^{k*}\alpha_k
\end{aligned}$$

and so

$$\begin{aligned} P_{S_N}(s) &= \sum_{j=0}^{\infty} s^j \sum_{k=0}^{\infty} p_j^{k*} \alpha_k \\ &= \sum_{k=0}^{\infty} \alpha_k \sum_{j=0}^{\infty} s^j p_j^{k*} \\ &= \sum_{k=0}^{\infty} \alpha_k P_{s_k}(s) = \sum_{k=0}^{\infty} \alpha_k \left(P_{X_1}(s)\right)^k \\ &= E\left[\left(P_{X_1}(s)\right)^N\right] \\ &= P_N\left(P_{X_1}(s)\right) \end{aligned}$$

**Example 1.11.** Suppose $N$ is $p(k;\lambda)$ and

$$X_1 = \begin{cases} 1 & \text{with prob. } p \\ 0 & \text{with prob. } 1-p \end{cases}$$

From our previous expression,

$$P_{s_N}(s) = P_N(P_{X_1}(s)) = \exp(\lambda(ps-p)) = \exp(\lambda p(s-1))$$

and $s_N$ is $p(k;\lambda p)$.

*Remark* 1.9. (Wald's identity) Note that

$$E[s_N] = \frac{d}{ds} P_N(P_{X_1}(s))\Big|_{s=1} = P_N'(P_{X_1}(1)) P_{X_1}'(1) = E[N]E[X_1]$$

## 1.3 Branching Processes

**Definition 1.4.** Let $\{Z_{n,j} : n \geq 1, j \geq 1\}$ be iid non-negative random variables having common probability mass functions $\{p_k\}$. Define $\{Z_n : n \geq 0\}$ by:

$$\begin{aligned} Z_0 &= 1 \\ Z_1 &= Z_{1,1} \\ Z_2 &= Z_{2,1} + Z_{2,2} + \ldots + Z_{2,Z_1} \\ &\vdots \\ Z_n &= Z_{n,1} + Z_{n,2} + \ldots + Z_{n,Z_{n-1}} \end{aligned}$$

If $Z_n = 0$ then $Z_{n+1} = 0$. This is a **branching process**.

*Remark* 1.10. Define $P_n(s) = E(s^{Z_n})$ and $P(s) = E(s^{Z_1}) = \sum_{k=0}^{\infty} p_k s^k$ and note that

$$\begin{aligned} P_0(s) &= s \\ P_1(s) &= P(s) = E(s^{Z_1}) = \sum_{k=0}^{\infty} p_k s^k \\ P_2(s) &= P_1(P(s)) = P(P(s)) \\ P_3(s) &= P_2(P(s)) = P(P(P(s))) = P(P_2(s)) \\ &\vdots \\ P_n(s) &= P_{n-1}(P(s)) = P(P_{n-1}(s)) \end{aligned}$$

**Example 1.12.** Suppose $Z_{n,j}$ is a Bernoulli random variable which is equal to 1 with probability $p$ and $0$ otherwise. Then

$P(s) = (1-p) + ps$ and

$$\begin{aligned} P_2(s) &= (1-p) + p(1-p) + p^2 s \\ P_3(s) &= (1-p) + p(1-p) + p^2(1-p) + p^3 s \\ &\vdots \\ P_n(s) &= (1-p) + p(1-p) + ... + (1-p)p^{n-1} + p^n s \\ &= \left[(1-p)\sum_{k=0}^{n-1} p^k\right] + p^n s \end{aligned}$$

**Example 1.13.** What is $E(Z_n)$? Suppose that $E(Z_1) = m$. Then

$$\begin{aligned} P_n'(s) &= P'(P_{n-1}(s))P_{n-1}'(s) \\ P_n'(1) &= mP_{n-1}'(1) \\ &= m^2 P_{n-2}'(1) \\ &\vdots \\ &= m^n \end{aligned}$$

*Remark* 1.11. Consider the event $\{\text{extinction}\} = \bigcup_{n=1}^{\infty}\{Z_n = 0\}$. Let $\Pi = P(\{\text{extinction}\}) = P\left(\bigcup_{n=1}^{\infty}\{Z_n = 0\}\right)$. Note that $\{Z_n = 0\} \subset \{Z_{n+1} = 0\}$. We have

$$\begin{aligned} \Pi = P\left(\bigcup_{n=1}^{\infty}\{Z_k = 0\}\right) &= \lim_{n\to\infty} P\left(\bigcup_{k=1}^{n}\{Z_k = 0\}\right) \\ &= \lim_{n\to\infty} P(Z_n = 0) \\ &= \lim_{n\to\infty} P_n(0) \end{aligned}$$

where $P_n(s) = E(s^{Z_n})$. This is a very difficult method of determining extinction probability.

*Remark* 1.12. Consider iid $\{Z_{n,j} : n \geq 1, j \geq 1\}$ having probability mass function $\{p_k\}$. Note that if

$$\begin{aligned} p_0 = 0 &\implies \Pi = 0 \\ p_0 = 1 &\implies \Pi = 1 \end{aligned}$$

We will now consider the case where $0 < p_0 < 1$.

**Theorem 1.1.** *If $m = E[Z_1] < 1$ then $\Pi = 1$. If $m > 1$ then $\Pi < 1$ and is the unique non-negative solution to the equation $s = P(s)$ which is less than 1.*

*Proof.* Let us first show that $\Pi$ is a solution of $s = P(s)$ and define $\Pi_n = P(Z_n = 0)$ where $\{\Pi_n\}$ is a non-decreasing sequence converging to $\Pi$. Recall that

$$P_{n+1}(s) = P(P_n(s))) \implies \Pi_{n+1} = P(\Pi_n) \text{ at } s = 0$$

and hence

$$\Pi = \lim_{n\to\infty} \Pi_{n+1} = \lim_{n\to\infty} P(\Pi_n) = P(\Pi)$$

Next we show that $\Pi$ is the smallest solution of $P(s) = s$ in $[0, 1]$. Suppose $q$ is some other solution to $P(s) = s$ with $0 \leq q \leq 1$. Note that

$$\begin{aligned} \Pi_1 &= P(0) \leq P(q) = q \\ \Pi_2 &= P(\Pi_1) \leq P(q) = q \\ &\vdots \\ \Pi_n &\leq q \end{aligned}$$

as $n \to \infty$ then $\Pi_n \to \Pi$ and $\Pi \leq q$. Finally note that $P(s)$ is convex since $P''(s) = \sum_{k=2}^{\infty} k(k-1)p_k s^{k-2} \geq 0$. Suppose

$m < 1 \implies P'(1) = E(Z_1) = m < 1$. If $P'(1) = m \leq 1$ then in a left neighbourhood of 1, $P(s)$ cannot be below the line $y = s$ and similarly if $P'(1) = m > 1$ in a left neighbourhood of 1, $P(s)$ must intersect $y = s$ at some point $0 < s < 1$ (see Resnick, p. 23). □

## 1.4 Continuity Theorem

Let $\{X_n : n \geq 0\}$ be non-negative integer valued random variables with

$$P(X_n = k) = p_k^{(n)}, P_n(s) = E(s^{X_n})$$

Then $X_n$ converges in distribution to $X_0$ if

$$\lim_{n\to\infty} p_k^{(n)} = p_k^{(0)}, \forall k = 0, 1, ...$$

**Theorem 1.2.** *Suppose for each $n = 1, 2, ...$ $\{p_k^{(n)} : k \geq 0\}$ is a probability mass function $\{0, 1, 2, ...\}$ so that*

$$p_k^{(n)} \geq 0, \sum_{k=0}^{\infty} p_k^{(n)} = 1$$

*Then there exists a sequence $\{p_k^{(0)} : k \geq 0\}$ such that $\lim_{n\to\infty} p_k^{(n)} = p_k^{(0)}$ for all $k = 0, 1, ...$ if and only if there exists a function $P_0(s)$, $0 < s < 1$ such that*

$$\lim_{n\to\infty} P_n(s) = P_0(s)$$

*Proof.* ( $\implies$ ) Suppose $p_k^{(n)} \to p_k^{(0)}$ and fix $s \in (0, 1)$, $\epsilon > 0$ and pick $m$ large enough such that

$$\sum_{i=m+1}^{\infty} s^i < \epsilon$$

Then observe that

$$\begin{aligned}
|P_n(s) - P_0(s)| &= \left|\sum_{k=0}^{\infty} p_k^{(n)} s^k - \sum_{k=0}^{\infty} p_k^{(0)} s^k\right| \\
&\leq \sum_{k=0}^{\infty} \left|p_k^{(n)} - p_k^{(0)}\right| s^k \\
&= \sum_{k=0}^{m} |p_k^{(n)} - p_k^{(0)}| s^k + \sum_{k=m+1}^{\infty} |p_k^{(n)} - p_k^{(0)}| s^k \\
&\leq \sum_{k=0}^{m} |p_k^{(n)} - p_k^{(0)}| s^k + \sum_{k=m+1}^{\infty} s^k \\
&\leq \sum_{k=0}^{m} |p_k^{(n)} - p_k^{(0)}| s^k + \epsilon
\end{aligned}$$

Hence,

$$\lim_{n\to\infty} |P_n(s) - P_0(s)| < \epsilon$$

and since $\epsilon$ was arbitrary, we are done.

( $\impliedby$ ) For a fixed $k$ let $\left\{p_k^{(n')}\right\}$ be a subsequence such that $\lim_{n\to\infty} p_k^{(n')}$ exists. Let $\left\{p_k^{(n'')}\right\}$ be another subsequence such that

$\lim_{n\to\infty} p_k^{(n'')}$ exists. Remark that

$$\lim_{n'\to\infty}\sum_{k=0}^{\infty} p_k^{(n')} s^k = \lim_{n'\to\infty} P_{n'}(s) = P_0(s)$$
$$\lim_{n''\to\infty}\sum_{k=0}^{\infty} p_k^{(n'')} s^k = \lim_{n''\to\infty} P_{n''}(s) = P_0(s)$$

Then the two subsequences have the same probability generating function. Since the probability generating function uniquely defines the probability mass function, all subsequences yield the same limit and hence $\lim_{n\to\infty} p_k^{(n)}$ exists. □

## 1.5 Random Walk

**Definition 1.5.** Let $\{X_n : n \geq 1\}$ be iid random variables (r.v.s) taking values $-1$ and $1$. with $P(X_1 = 1) = p$ and $P(X_1 = -1)$. Let

$$S_0 = 0, S_1 = X_1, ..., S_n = \sum_{k=1}^{n} X_k$$

Then $\{S_n : n \geq 0\}$ is called the **simple random walk**.

*Remark* 1.13. Define $N = \inf\{n \geq 1 : S_n = 1\}$ and $\phi_n = P(N = n)$ with $\phi_0 = 0, \phi_1 = p$. For $n \geq 2$, suppose we have 1 step of $0 \to -1$, it takes $j$ steps to get $-1 \to 0$, and $k$ steps to get $0 \to 1$. Then we should have $1 + j + k = n$ with

$$\phi_n = \sum_{j=1}^{n-2}(1-p)\phi_j\phi_{n-j-1}$$

with more details below:

$$\{N = n\} = \bigcup_{j=1}^{n-2}\{X_1 = -1\} \cap A_j \cap B_{n-j-1}$$
$$A_j = \left\{\inf\left\{n : \sum_{i=1}^{n} X_{i+1} = 1\right\} = j\right\}$$
$$B_{n-j-1} = \left\{\inf\left\{n : \sum_{i=1}^{n} X_{i+j+1} = 1\right\} = n-j-1\right\}$$

Since $A_j$ is independent of $B_{n-j-1}$ then

$$P(N = n) = \sum_{j=1}^{n-2}(1-p)P(A_j)P(B_{n-j-1})$$
$$= \sum_{j=1}^{n-2}(1-p)\phi_j\phi_{n-j-1}$$

Now define $\Phi(s) = \sum_{n=0}^{\infty} s^n \phi_n$ and note that

$$\begin{aligned}
\Phi(s) - ps = \sum_{n=2}^{\infty} \phi_n s^n &= \sum_{n=2}^{\infty} s^n \sum_{j=1}^{n-2} (1-p)\phi_j \phi_{n-j-1} \\
&= (1-p) \sum_{n=2}^{\infty} s^n \sum_{j=0}^{n-2} (1-p)\phi_j \phi_{n-j-1} \\
&= (1-p) \sum_{j=0}^{\infty} \sum_{n=j+2}^{\infty} s^n \phi_j \phi_{n-j-1} \\
&= (1-p) \sum_{j=0}^{\infty} \sum_{n=j+2}^{\infty} s^j \phi_j s^{n-j} \phi_{n-j-1} \\
&= (1-p)s \sum_{j=0}^{\infty} s^j \phi_j \sum_{n=j+2}^{\infty} s^{n-j-1} \phi_{n-j-1} \\
&= (1-p)s\Phi^2(s)
\end{aligned}$$

and we have the following quadratic: $(1-p)s\Phi^2(s) - \Phi(s) + ps = 0$ with the solution

$$\Phi(s) = \frac{1 \pm \sqrt{1-4p(1-p)s^2}}{2(1-p)s}$$

Note that

$$\Phi(0) = \lim_{s \to 0} \frac{1 + \sqrt{1-4p(1-p)s^2}}{2(1-p)s} = \infty$$

so it must be the case that

$$\Phi(s) = \frac{1 - \sqrt{1-4p(1-p)s^2}}{2(1-p)s}$$

*Remark* 1.14. With our new function, we can get

$$P(N < \infty) = \Phi(1) = \frac{1 - \sqrt{1-4p(1-p)}}{2(1-p)} = \frac{1 - |2p-1|}{2(1-p)}$$

If $p \le 1/2$ then

$$P(N < \infty) = \frac{1-1+2p}{2(1-p)} = \frac{p}{1-p} \implies P(N = \infty) = \frac{1-2p}{1-p} \implies E(N) = \infty$$

But if $p \ge 1/2$ then

$$P(N < \infty) = \frac{2-2p}{2(1-p)} = 1$$

Let's calculate $E(N)$ when $p \ge 1/2$. First note that

$$\Phi'(1) = \frac{2p}{|2p-1|} - \frac{1-|2p-1|}{2(1-p)}$$

and hence

$$E(N) = \begin{cases} \infty & p = \frac{1}{2} \\ \frac{1}{2p-1} & p > \frac{1}{2} \end{cases}$$

*Remark* 1.15. Let $N_0 = \inf\{n \ge 1 : S_n = 0\}$ and $f_n = P(N = n)$, $f_0 = 0$ with observation that only $f_{2n} = P(N = 2n) > 0$ for $n = 1, 2, \ldots$. Let

$$F(s) = \sum_{n=0}^{\infty} s^{2n} f_{2n}$$

If $X_1 = -1$ then $N_0 = 1 + \inf\{n : \sum_{i=1}^{n} X_{i+1} = 1\} = 1 + N^+$ and if $X_1 = 1$ then $N_0 = 1 + \inf\{n : \sum_{i=1}^{n} X_{i+1} = -1\} = 1 + N^-$

with the remark that $P(N^+ = n) = \phi_n$. Now

$$\begin{aligned}
F(s) &= E\left[s^{N_0}\right] = E\left[s^{N_0}1\{X_1=-1\}\right] + E\left[s^{N_0}1\{X_1=1\}\right] \\
&= E\left[s^{1+N^+}1\{X_1=-1\}\right] + E\left[s^{1+N^-}1\{X_1=1\}\right] \\
&= s(1-p)E\left[s^{N^+}\right] + spE\left[s^{N^-}\right] \\
&= s(1-p)\Phi(s) + spE\left[s^{N^-}\right]
\end{aligned}$$

Now,

$$\begin{aligned}
N^- = \inf\left\{n : \sum_{i=1}^{n} x_{i+1} = -1\right\} &\stackrel{d}{=} \inf\left\{n : \sum_{i=1}^{n} x_i = -1\right\} \\
&= \inf\left\{n : \sum_{i=1}^{n} (-x_i) = 1\right\}
\end{aligned}$$

and hence $P(-X_1 = 1) = 1-p, P(-X_1 = -1) = p$ and

$$E[s^{N^-}] = \frac{1 - s\sqrt{1-4p(1-p)s^2}}{2ps}$$

with the final result

$$\begin{aligned}
F(s) &= s(1-p)\frac{1-\sqrt{1-4p(1-p)s^2}}{2(1-p)s} + sp\frac{1-s\sqrt{1-4p(1-p)s^2}}{2ps} \\
&= 1 - \sqrt{1-4p(1-p)s^2}
\end{aligned}$$

*Remark* 1.16. Let's calculate

$$\begin{aligned}
P(N_0 < \infty) = F(s) = 1 - \sqrt{(1-2p)^2} &= 1 - |1-2p| \\
&= \begin{cases} 1 & p = \frac{1}{2} \\ 2(1-p) & p > \frac{1}{2} \\ 2p & p < \frac{1}{2} \end{cases}
\end{aligned}$$

So $E[N_0] = \infty$ for $p \neq 1/2$. However, also note that if $p = 1/2$ then

$$E[N_0] = F'(1) = \lim_{s\to 1} F'(s) = \lim_{s\to 1} \frac{s}{\sqrt{1-s^2}} = \infty$$

# 2 Discrete Time Markov Chains

*Remark* 2.1. Let $P(X = k) = a_k$ for $k = 0, 1, ...$with $\sum_{k=0}^{\infty} a_i = 1$. Suppose $U$ is a uniform random variable in $(0,1)$ and define

$$Y = \sum_{k=0}^{\infty} k1\left(\sum_{i=0}^{k-1} a_i, \sum_{i=1}^{k} a_i\right)(U)$$

where $1(a,b)(U)$ is 1 if $a \leq U \leq b$ and 0 otherwise. Then $X$ and $Y$ have the same probability mass function. So $Y = k$ if and only if $U \in \left(\sum_{i=0}^{k-1} a_i, \sum_{i=1}^{k} a_i\right)$.

**Definition 2.1.** Given $S = \{0, 1, 2, ...\}$ with $a_k = P(X_0 = k)$ and define $\boldsymbol{P} = \{p_{ij} : i \geq 0, j \geq 0\}$ which we call the **probability transition matrix**. Define

$$X_0 = \sum_{k=0}^{\infty} k1\left(\sum_{i=0}^{k-1} a_i, \sum_{i=1}^{k} a_i\right)(U_0)$$

and $f(i,u)$ on $S \times [0,1]$ as

$$f(i,u) = \sum_{k=0}^{\infty} 1\left(\sum_{j=0}^{k-1} p_{ij}, \sum_{j=0}^{k} p_{ij}\right)(u)$$

where $f(i,u) = k$ if and only if $u \in \left(\sum_{j=0}^{k-1} p_{ij}, \sum_{j=0}^{k} p_{ij}\right)$. Now define $X_{n+1} = f(X_n, U_{n+1})$ where $X_n$ depends on $X_{n-1}, U_0, U_1, ..., U_n$.

Here are some properties:

(1) $P(X_0 = k) = a_k$ and

$$\begin{aligned} P(X_{n+1} = j | X_n = i) &= P\left(f(X_n, U_{n+1}) = j | X_n = i\right) \\ &= P\left(f(i, U_{n+1}) = j\right) \\ &= p_{ij} \end{aligned}$$

(2) [Markov Property] We can see from (1) that

$$\begin{aligned} P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, ..., X_n = i) &= P(f(X_n, U_{n+1}) | X_0 = i_0, X_1 = i_1, ..., X_n = i) \\ &= P(f(i, U_{n+1}) = j) \\ &= p_{ij} \end{aligned}$$

(3) A application of the above is

$$\begin{aligned} P(X_{n+1} = k_1, X_{n+2} = k_2, ..., X_{n+m} = k_m | X_0 = i_0, ..., X_n = i) &= P(X_{n+1} = k_1, X_{n+2} = k_2, ..., X_{n+m} = k_m | X_n = i) \\ &= P(X_1 = k_1, X_2 = k_2, ..., X_m = k_m | X_0 = i_0, ..., X_n = i) \end{aligned}$$

**Definition 2.2.** Any stochastic process $\{X_n : n \geq 0\}$ satisfying $P(X_{n+1} = j | X_n = i) = p_{ij}$ and $P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, ...., X_n = i) = p_{ij}$ is a called a **Markov chain** with initial distribution $\{a_k\}$ and probability transition matrix $\boldsymbol{P}$.

**Proposition 2.1.** *Given a Markov chain, the finite dimensional distributions are given of the form*

$$P(X_0 = i_0, ..., X_k = i_k) = a_{i_0} p_{i_0 i_1} ... p_{i_{k-1} i_k}$$

*Proof.* (1) Suppose that

$$P(X_{i_0} = i_0, ..., X_j = i_j) > 0$$

for all $j = 0, ..., k-1$. Then

$$\begin{aligned} P(X_0 = i_0, ..., X_k = i_k) &= P(X_k = i_k | X_0 = i_0, ..., X_{k-1} = i_{k-1}) P(X_0 = i_0, ..., X_{k-1} = i_{k-1}) \\ &= p_{i_{k-1} i_k} P(X_{k-1} = i_{k-1} | X_0 = i_0, ..., X_{k-1} = i_{k-2}) P(X_0 = i_0, ..., X_{k-1} = i_{k-2}) \\ &= p_{i_{k-1} i_k} p_{i_{k-2} i_{k-1}} ... p_{i_0 i_1} a_{i_0} \end{aligned}$$

Now suppose that there exists a $j$ such that

$$P(X_{i_0} = i_0, ..., X_j = i_j) = 0$$

and let

$$j^* = \inf\{j \geq 0 : P(X_0 = i_0, ..., X_j = i_j) = 0\}$$

If $j^* = 0$, then $P(X_0 = i_0) = 0$ and the result holds trivially. If $j^* > 0$ then $P(X_{i_0} = i_0, ..., X_{j^*-1} = i_{j^*-1}) > 0$ and hence

$$\begin{aligned} P(X_{i_0} = i_0, ..., X_{j^*} = i_{j^*}) &= P(X_{j^*} = i_{j^*} | X_0 = i_0, ..., X_{j^*-1} = i_{j^*-1}) P(X_0 = i_0, ..., X_{j^*-1} = i_{j^*-1}) \\ &= p_{i_{j^*-1} i_{j^*}} \times 0 \\ &= 0 \end{aligned}$$

(2) Conversely, given a density $\{a_k\}$, a transition matrix $\boldsymbol{P}$, and a process $\{X_n\}$ whose finite dimensional distribution is given as

$$P(X_0 = i_0, ..., X_k = i_k) = a_{i_0} p_{i_0 i_1} ... p_{i_{k-1} i_k}$$

then $\{X_n\}$ is a Markov chain with

$$\begin{aligned}
P(X_0 = k) &= a_k \\
P(X_{n+1} = j | X_n = i) &= p_{ij} \\
P(X_{n+1} = j | X_0 = i_0, ..., X_n = i) &= p_{ij} \\
P(X_{n+1} = j | X_0 = i_0, ..., X_n = i) &= \frac{P(X_{n+1} = j, X_n = i, ..., X_0 = i_0)}{P(X_n = i, ..., X_0 = i_0)} = p_{ij}
\end{aligned}$$

□

**Example 2.1.** (Branching process) The branching process $\{Z_n\}$ has

$$\begin{aligned}
P(Z_n = i_n | Z_0 = i_0, ..., Z_{n-1} = i_{n-1}) &= P\left(\sum_{j=1}^{i_{n-1}} Z_{n,j} = i_n \middle| Z_0 = i_0, ..., Z_{n-1} = i_{n-1}\right) \\
&= P\left(\sum_{j=1}^{i_{n-1}} Z_{n,j} = i_n\right)
\end{aligned}$$

and since

$$P(Z_{n+1} = j | Z_n = i) = P\left(\sum_{k=1}^{i} Z_{n,k} = j\right) = p_j^{*i}$$

the branching process is Markov and computable.

**Example 2.2.** (Random walk) Let $\{X_n\}$ be iid random variables with $P(X_n = k) = a_k$ and define $S_0 = 0, S_n = \sum_{i=1}^n X_i$. Then

$$\begin{aligned}
P(S_{n+1} = i_{n+1} | S_0 = 0, ..., S_n = i_n) &= P(S_n + X_{n+1} = i_{n+1} | S_0 = 0, ..., S_n = i_n) \\
&= P(i_n + X_{n+1} = i_n)
\end{aligned}$$

and

$$p = P(X_{n+1} = i_{n+1} - i_n) = a_{i_{n+1} - i_n}$$

**Example 2.3.** (Inventory model) Let $I(t)$ denote the inventory level at time $t$. Suppose the inventory level is checked at fixed times $T_0, T_1, T_2, ....$ Define $X_n = I(T_n)$. If $X_n \leq s$, purchase enough units to bring the inventory level to $S$. Otherwise do not purchase any new items. Assume that new units are replenished in a negligible amount of time. Let $D_n$ be the demand during $[T_{n-1}, T_n]$ and assume $\{D_n, n \geq 0\}$ is a sequence of independent and identically distributed random variables and independent of $X_0$. Suppose $X_0 \leq S$ and no backlogs are allowed. Then,

$$X_{n+1} = \begin{cases} \max(X_n - D_{n+1}, 0) & X_n > s \\ \max(S - D_{n+1}, 0) & X_n \leq s \end{cases}$$

with state space $\{0, 1, ..., S\}$.

**Example 2.4.** (Discrete time queue)

(1) Consider a queuing model where $T_0, T_1, T_2, ...$ denote the departure times from the system. Let $X(t)$ be the number of customers at time $t$ and $X_n = X(T_n^+)$ where $T_n^+$ is the time right after the $n^{th}$ departure. Let $A_n$ denote the number of arrivals in the time interval $[T_{n-1}, T_n)$. Then

$$X_{n+1} = \max(X_n + A_{n+1} - 1, 0)$$

If $P(A_1 = k) = a_k$ then this is a discrete time Markov process with transition matrix

$$\boldsymbol{P}_{ij} = \begin{cases} 0 & i - j \geq 1 \\ a_0 + a_1 & i = j = 0 \\ a_{j-i+1} & o/w \end{cases}$$

(2) Let $T_0, T_1, T_2, ...$ denote the times that customers arrive at the system. Let $X_n = X(T_n^-)$ where $T_n^-$ is the time right after the $n^{th}$ arrival and $S_{n+1}$ be the number of service completions in the time interval $[T_n, T_{n+1})$ with state space $\{0, 1, 2, ...\}$. Then

$$X_{n+1} = \max(X_n - S_{n+1} + 1, 0)$$

If $P(S_1 = k) = b_k$ then this is a discrete time Markov process with transition matrix

$$\boldsymbol{P}_{ij} = \begin{cases} \sum_{k=i+1}^{\infty} b_k & j = 1 \\ 0 & j - i \geq 2 \\ b_{i-j+1} & o/w \end{cases}$$

**Proposition 2.2.** *Using the notation* $p_{ij}^{(2)} = (\boldsymbol{P}^2)_{ij} = \sum_k p_{ik} p_{kj}$ *and* $p_{ij}^{(n)} = \sum_k p_{ik} p_{kj}^{(n-1)} = \sum_k p_{ik}^{(n-1)} p_{kj}$*, we have for all* $n \geq 0$ *and* $i, j \in S$

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i)$$

*Proof.* Clearly it holds for $n = 0, 1$. Now suppose it holds for $0, 1, ..., n$. Then

$$\begin{aligned} P(X_{n+1} = j | X_0 = i) &= \sum_k P(X_{n+1} = j, X_1 = k | X_0 = i) \\ &= \sum_k P(X_{n+1} = j | X_1 = k, X_0 = i) P(X_1 = k | X_0 = i) \\ &= \sum_k P(X_{n+1} = j | X_1 = k) P(X_1 = k | X_0 = i) \\ &= \sum_k P(X_n = j | X_0 = k) P(X_1 = k | X_0 = i) \\ &= \sum_k p_{kj}^{(n)} p_{ik} \\ &= \sum_k p_{ik} p_{kj}^{(n)} = \sum_k p_{ik}^{(n)} p_{kj} \end{aligned}$$

□

*Notation* 1. We call the equation

$$p_{ij}^{(n+m)} = \sum_k p_{ik}^{(n)} p_{kj}^{(m)}$$

the **Chapman-Komolgorov equation.**

**Corollary 2.1.** $P(X_n = j) = \sum_i a_i p_{ij}^{(n)}$

*Proof.* Immediate from

$$P(X_n = j) = \sum_i P(X_n = j | X_0 = i) a_i = \sum_i p_{ij}^{(n)} a_i$$

□

*Notation* 2. From the book we will denote $P(X_n = j) = a_j^{(n)}$.

## 2.1 State Space Decomposition

Let $\{X_n : n \geq 0\}$ be a Markov chain with state space $S$. Set $B \subset S$ and $\tau_B = \inf\{n \geq 0 : X_n \in B\}$ which we call the hitting time of $B$. We use $\tau_j = \tau_{\{j\}}$.

**Definition 2.3.** For $i, j \in S$ we say state $j$ is **accessible** from state $i$ if

$$P(\tau_j < \infty | X_0 = i) > 0$$

and we denote it as $i \to j$. Obviously $i \to i$.

**Proposition 2.3.** *For $i \neq j$ we have $i \to j$ if and only if there exists $n > 0$ such that $p_{ij}^{(n)} > 0$. That is, $P(X_n = j | X_0 = i) > 0$.*

*Proof.* Suppose that there exists $n$ such that $p_{ij}^{(n)} > 0$ and note that

$$\{X_n = j\} \subseteq \{\tau_j \leq n\} \subseteq \{\tau_j < \infty\} \implies 0 < P(X_n = j|X_0 = i) \subseteq P(\tau_j \leq n|X_0 = i) \subseteq P(\tau_j < \infty|X_0 = i)$$

Now suppose that $P(\tau_j < \infty|X_0 = i) > 0$ and assume that $p_{ij}^{(n)} = 0$ for all $n$. Then

$$\begin{aligned} P(\tau_j < \infty|X_0 = i) &= \lim_{n\to\infty} P(\tau_j \leq n|X_0 = i) \\ &= \lim_{n\to\infty} P\left(\bigcup_{k=0}^{n} \{X_k = j\}|X_0 = i\right) \\ &\leq \limsup_{n\to\infty} \sum_{k=0}^{n} P(X_k = j|X_0 = i) = 0 \end{aligned}$$

which is a contradiction. □

**Definition 2.4.** States $i$ and $j$ **communicate** $i \leftrightarrow j$ if they are accessible from each other (i.e. $i \to j$ and $j \to i$). Communication is an equivalence class as follows

(1) $i \leftrightarrow i$ (reflexive)

(2) $i \leftrightarrow j$ if and only if $j \leftrightarrow i$ (symmetric)

(3)) $i \leftrightarrow j$ and $j \leftrightarrow j$ then $i \leftrightarrow k$ (transitive)

(1) and (2) are obvious. For (3) suppose $n$ and $m$ are such that $p_{ij}^{(n)} > 0$ and $p_{jk}^{(m)} > 0$. Then $p_{ik}^{(n+m)} = \sum_l p_{il}^{(n)} p_{lk}^{(m)} > 0$ and we are done.

*Remark* 2.2. We can then partition the state space into equivalence classes $C_0, C_1, ...$ such that

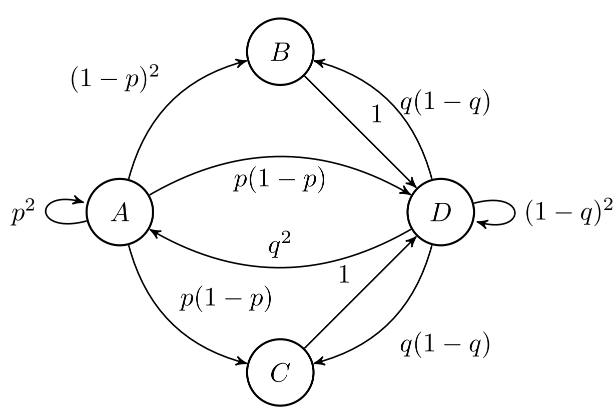
$$C_i \cap C_j = \emptyset, \bigcup_i C_i = S$$

**Example 2.5.** Consider a Markov chain with state space $\{0, 1, 2, 3\}$ and

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with equivalence classes $\{0\}, \{1, 2\}, \{3\}$.

*Notation* 3. Here is one way to represent Markov chains (with a Markov probability transition diagram):



**Example 2.6.** Now consider a Markov chain with $S = \{1, 2, 3, 4\}$ and

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

and $\{1, 2\}, \{3, 4\}$ are equivalence classes.

**Example 2.7.** $S = \{0, 1, 2, ...\}$ is an equivalence class for $\mathbb{P}(S_1 = j) = a_j$ (note we may use $\mathbb{P}$ and $P$ interchangeably for "probability of") and

$$P = \begin{pmatrix} \sum_{i=1}^{\infty} a_i & a_0 & 0 & 0 & \cdots & \\ \sum_{i=1}^{\infty} a_i & a_1 & a_0 & 0 & \cdots & \\ \sum_{i=1}^{\infty} a_i & a_2 & a_1 & a_0 & 0 & \cdots \\ \vdots & & & & & \end{pmatrix}$$

This is an example of an irreducible Markov chain.

**Definition 2.5.** A Markov chain is **irreducible** if the state space consists of only one equivalence class. This means that $i \leftrightarrow j$ for all $i, j \in S$.

**Definition 2.6.** A set of states $C \subset S$ is **closed** if for any $i \in C$ we have $P(\tau_{C^c} = \infty | X_0 = i) = 1$. If a singleton is closed then it is called an **absorbing state**.

**Proposition 2.4.** *(i) $C$ is closed if and only if for all $i \in C$ and $j \in C^c$ we have $p_{ij} = 0$.*

*(ii) $j$ is absorbing if and only if $p_{jj} = 1$.*

*Proof.* (i) ($\implies$)Suppose that $P(\tau_{C^c} = \infty | X_0 = i) = 1$. Then we know that there exists no $n$ such that $p_{ij}^{(n)} > 0$ for $j \in C^c$ and then clearly $p_{ij} = 0$ for $j \in C^c$.

($\impliedby$) Conversely suppose that $p_{ij} = 0$ for all $j \in C^c$. Then,

$$P(\tau_{C^c} = 1 | X_0 = i) = \sum_{j \in C^c} p_{ij} = 0$$

and

$$\begin{aligned} P(\tau_{C^c} \leq 2 | X_0 = i) &= P(\tau_{C^c} = 1 | X_0 = i) + P(\tau_{C^c} = 2 | X_0 = i) \\ &= 0 + P(X_1 \in C, X_2 \in C^c | X_0 = i) \\ &= \sum_{j \in C^c} \sum_{k \in C} p_{ik} p_{kj} = 0 \end{aligned}$$

Continuing in this manner, we have $P(\tau_{C^c} \leq n | X_0 = i) = 0$ and thus $\lim_{n \to \infty} P(\tau_{C^c} \leq n | X_0 = i) = 0$.

(ii) This is obvious. □

**Example 2.8.** Consider

$$X_{n+1} = \begin{cases} \max(X_n - D_{n+1}, 0) & X_n > s \\ \max(S - D_{n+1}, 0) & X_n \leq s \end{cases}$$

with $X_0 < S$ and $P(D_1 = k) = p_k$.

$$P = \begin{pmatrix}
\underbrace{\sum_{k=S}^{\infty} p_k}_{P_{11}} & p_{S-1} & p_{S-2} & \cdots & & & & & & \underbrace{p_0}_{P_{1S}} \\
\vdots & & & & & & & & & \vdots \\
\underbrace{\sum_{k=S}^{\infty} p_k}_{P_{s1}} & p_{S-1} & p_{S-2} & \cdots & & & & & & p_0 \\
\underbrace{\sum_{k=s+1}^{\infty} p_k}_{P_{(s+1)1}} & p_S & p_{S-1} & \cdots & \underbrace{p_1}_{P_{(s+1)s}} & p_0 & 0 & & \cdots & 0 \\
\underbrace{\sum_{k=s+2}^{\infty} p_k}_{P_{(s+2)1}} & p_S & p_{S-1} & \cdots & \underbrace{p_2}_{P_{(s+1)s}} & p_1 & p_0 & 0 & \cdots & 0 \\
\vdots & & & & & & & & & \vdots \\
\underbrace{\sum_{k=S}^{\infty} p_k}_{P_{S1}} & p_{S-1} & \cdots & & & & & & & \underbrace{p_0}_{P_{SS}}
\end{pmatrix}$$

Note that since $0 \to i$ and $i \to 0$ for any $i \in S$ then this system is irreducible.

**Example 2.9.** Suppose that $P(X_0 = i) = 1$ and define $\tau_i(0) = 0, \tau_i(1) = \inf\{m \geq 1 : X_m = i\}$. Suppose that $\tau_i(1) < \infty$ and define $\tau_i(2) = \inf\{m > \tau_i(1) : X_m = i\}$. Continuing in this manner, assuming that $\tau_i(n) < \infty$, then we define

$$\tau_i(n+1) = \inf\{m > \tau_i(n) : X_m = i\}$$

Let $\alpha_0 = 0, \alpha_1 = \tau_i(1), \alpha_2 = \tau_i(2) - \tau_i(1), ..., \alpha_n = \tau_i(n) - \tau_i(n-1)$ and define

$$\begin{aligned}
\varepsilon_1 &= (\alpha_1, X_1, X_2, ..., X_{\tau_i(1)}) \\
\varepsilon_2 &= (\alpha_2, X_{\tau_i(1)+1}, X_{\tau_i(1)+2}, ..., X_{\tau_i(2)}) \\
&\vdots \\
\varepsilon_n &= (\alpha_n, X_{\tau_i(n-1)+1}, X_{\tau_i(n-1)+2}, ..., X_{\tau_i(n)})
\end{aligned}$$

on $\tau_i(1) < \infty, \tau_i(2) < \infty, ..., \tau_i(n) < \infty$.

**Proposition 2.5.** *Suppose that $X_0 = i$. Then we have $\varepsilon_1, \varepsilon_2, ..., \varepsilon_k$ are iid with respect to the probability measure*

$$P(\cdot | \tau_i(1) < \infty, ...., \tau_i(k) < \infty)$$

*Proof.* Consider

$$P(\varepsilon_1 = (k, i_1, i_2, ..., i_k), \varepsilon_2 = (l, j_1, j_2, ..., j_k), \tau_i(1) < \infty, \tau_i(2) < \infty)$$

We need $i_k = i, j_k = i$ and furthermore $i_1 \neq i, ..., i_{k-1} \neq i$ and $j_1 \neq i, ..., j_{l-1} = i$. So,

$$\begin{aligned}
&P(\varepsilon_1 = (k, i_1, i_2, ..., i_k), \varepsilon_2 = (l, j_1, j_2, ..., j_k), \tau_i(1) < \infty, \tau_i(2) < \infty) \\
&= P(X_1 = i_1, X_2 = i_2, ..., X_{k-1} = i_{k-1}, X_k = i, X_{k+1} = j_1, X_{k+2} = j_2, ...., X_{k+l-1} = j_{l-1}, X_{k+l} = i) \\
&= P(X_{k+1} = j_1, X_{k+2} = j_2, ...., X_{k+l} = i | X_1 = i_1, X_2 = i_2, ..., X_k = i) P(X_1 = i_1, X_2 = i_2, ..., X_k = i) \\
&= P(X_1 = j_1, X_2 = j_2, ...., X_l = i) P(X_1 = i_1, X_2 = i_2, ..., X_k = i) \\
&= P(X_1 = j_1, X_2 = j_2, ...., X_{l-1} = j_{l-1}, \tau_i(2) = l) P(X_1 = i_1, X_2 = i_2, ..., X_{k-1} = i_{k-1}, \tau_i(1) = k)
\end{aligned}$$

Summing over the margins that are not $\tau_i(1) = l, \tau_i(1) = k$ on both sides of the equation (wrt $j_1, j_2, ..., j_{l-1}, i_1, i_2, ..., i_{k-2}$), we get

$$P(\tau_i(1) = l)P(\tau_i(1) = k) = P(\tau_i(1) = k, \tau_i(2) = l)$$

which implies

$$\begin{aligned} P(\alpha_2 = l)P(\alpha_1 = k) &= P(\alpha_1 = k, \alpha_2 = l) \\ \implies P(\tau_1 < \infty)P(\tau < \infty) &= P(\tau_1 < \infty, \tau_2 < \infty) \end{aligned}$$

This process may be generalized for not just pairwise $\varepsilon_i$ but any arbitrary group of $\varepsilon_i$ and so we are done. □

**Corollary 2.2.** *Suppose initial state $j \neq i$. Then still have $\varepsilon_1, ..., \varepsilon_k$ with respect to*

$$P(\cdot|\tau_1 < \infty, ..., \tau_k < \infty)$$

*Note that $\varepsilon_1$ will no longer have the same distribution as $\varepsilon_2$.*

**Definition 2.7.** State $i$ is **recurrent** if the chain returns to $i$ in a finite number of steps. Otherwise it is **transient**. That is,

- State $i$ is recurrent if $P(\tau_i(1) < \infty | X_0 = i) = 1$
- State $i$ is transient if $P(\tau_i(1) < \infty | X_0 = i) < 1 \implies P(\tau_i(1) = \infty | X_0 = i) > 0$

A recurrent state is **positive recurrent** if $E[\tau_i(1)|X_0 = i] < \infty$. Otherwise if $E[\tau_i(1)|X_0 = i] = \infty$ then a recurrent state is **null recurrent**.

**Definition 2.8.** For $n \geq 1$ define

$$\begin{aligned} f_{jk}^{(0)} &= 0 \\ f_{jk}^{(n)} &= P(\tau_k(1) = n | X_0 = j) \\ f_{jk} &= \sum_{n=0}^{\infty} f_{jk}^{(n)} = P(\tau_k(1) < \infty | X_0 = j) \end{aligned}$$

Therefore, a state $i$ is recurrent if and only if $f_{ii} = 1$ and a recurrent state $i$ is positive recurrent if and only if

$$E[\tau_i(1)|X_0 = i] = \sum_{n=0}^{\infty} n f_{ii}^{(n)} < \infty$$

*Remark* 2.3. Define $F_{ij}(s) = \sum_{n=0}^{\infty} s^n f_{ij}^{(n)}$ and $P_{ij}(s) = \sum_{n=0}^{\infty} s^n p_{ij}^{(n)}$

**Proposition 2.6.** *a) We have for $i \in S$*

$$p_{ii}^{(n)} = \sum_{k=0}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)}, \forall n \geq 1$$

*and for $0 < s < 1$ we have*

$$P_{ii}(s) = \frac{1}{1 - F_{ii}(s)}$$

*b) We have for $i \neq j$*

$$P_{ij}^{(n)} = \sum_{k=0}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}, \forall n \geq 0$$

*and for $0 < s < 1$ we have*

$$P_{ij}(s) = F_{ij}(s)P_{jj}(s)$$

*Proof.* a) Remark that

$$\begin{aligned} P(X_n = i|X_0 = i) &= \sum_{k=1}^{n} P(X_n = i, \tau_i(1) = k|X_0 = i) \\ &= \sum_{k=1}^{n} P(X_{\tau_i(1)+n-k} = i, \tau_i(1) = k|X_0 = i) \\ &= \sum_{k=1}^{n} P(\tau_i(1) = k|X_0 = i)P(X_{n-k} = i|X_0 = i) \\ p_{ii}^{(n)} &= \sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \end{aligned}$$

Now with this result, the second part can be written as

$$\begin{aligned} \sum_{n=1}^{\infty} s^n p_{ii}^{(n)} &= \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \\ P_{ii}(s) - 1 &= \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \\ &= \sum_{k=1}^{\infty} s^k f_{ii}^{(k)} \sum_{n=k}^{\infty} s^{n-k} p_{ii}^{(n-k)} \\ &= F_{ii}(s)P_{ii}(s) \end{aligned}$$

and so $P_{ii}(s) - 1 = F_{ii}(s)P_{ii}(s) \implies F_{ii}(s) = 1/(1 - P_{ii}(s))$.

b) By direct evaluation,

$$\begin{aligned} p_{ij}^{(n)} &= P(X_n = j|X_0 = i) \\ &= \sum_{k=0}^{n} P(\tau_i(j) = k|X_0 = i)P(X_{n-k} = j|X_0 = j) \\ &= \sum_{k=0}^{n} f_{ij}^{(n)} p_{jj}^{(n-k)} \end{aligned}$$

and so

$$\sum_{n=0}^{\infty} s^n p_{ij}^{(n)} = \sum_{n=0}^{\infty} s^n \sum_{k=0}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \implies P_{ij}(s) = F_{ij}(s)P_{jj}(s)$$

□

**Corollary 2.3.** *A state $i$ is recurrent if and only if $f_{ii} = 1$ if and only if $P_{ii}(1) = \sum p_{ii}^{(n)} = \infty$. Thus $i$ is transient if and only if $f_{ii} < 1$ if and only $\sum p_{ii}^{(n)} < \infty$.*

*Remark* 2.4. Define $N_j = \sum_{n=1}^{\infty} 1(X_n = j)$ which denotes the number of visits to state $j$. Then

$$\begin{aligned} E[N_j|X_0 = i] &= E\left[\sum_{n=1}^{\infty} 1(X_n = 1)|X_0 = i\right] \\ &= \sum_{n=1}^{\infty} E\left[1(X_n = 1)|X_0 = i\right] \\ &= \sum_{n=1}^{\infty} P(X_n = j|X_0 = i) \\ E[N_j|X_0 = i] &= \sum_{n=1}^{\infty} p_{ij}^{(n)} \end{aligned}$$

That is, state $i$ is recurrent if and only if $E[N_i|X_0 = i] =?$.

**Proposition 2.7.** *(i) We have for $i, j \in S$ and non-negative integer $k$*

$$P(N_j = k|X_0 = i) = \begin{cases} 1 - f_{ii} & k = 0 \\ f_{ij} f_{jj}^{k-1}(1 - f_{jj}) & k \geq 1 \end{cases}$$

*(ii) If $j$ is transient, then for all states $i$*

$$P(N_j < \infty|X_0 = i) = 1$$

*and $E[N_j|X_0 = i] = f_{ij}/(1 - f_{jj})$ and $P(N_j = k|X_0 = j) = (1 - f_{jj})f_{jj}^k$.*

*(iii) If $j$ is recurrent then $P(N_j = \infty|X_0 = j) = 1$.*

*Proof.* (i) We first calculate

$$P(N_j \geq 1|X_0 = i) = P(\tau_j(1) < \infty|X_0 = i) = f_{ij}$$

and similarly,

$$\begin{aligned} P(N_j \geq k|X_0 = i) &= P(\tau_j(k) < \infty|X_0 = i) \\ &= P(\tau_j(1) < \infty, \tau_j(2) < \infty, ..., \tau_j(k) < \infty|X_0 = i) \\ &= P(\tau_j(1) < \infty|X_0 = i)\,[P(\tau_j(1) < \infty|X_0 = k)]^{k-1} \\ P(N_j \geq k|X_0 = i) &= f_{ij} f_{jj}^{k-1} \end{aligned}$$

Hence

$$\begin{aligned} P(N_j = k|X_0 = i) &= P(N_j \geq k|X_0 = i) - P(N_j \geq k + 1|X_0 = i) \\ &= f_{ij} f_{jj}^{k-1} - f_{ij} f_{jj}^k \\ P(N_j = k|X_0 = i) &= f_{ij} f_{jj}^{k-1}(1 - f_{jj}) \end{aligned}$$

(ii) We can directly calculate

$$\begin{aligned} P(N_j = \infty|X_0 = i) &= \lim_{k\to\infty} P(N_j \geq k|X_0 = i) \\ &= \lim_{k\to\infty} f_{ij} f_{jj}^k = 0 \end{aligned}$$

and

$$\begin{aligned} E[N_j|X_0 = i] &= \sum_{k=0}^{\infty} P(N_j > k|X_0 = i) \\ &= \sum_{k=0}^{\infty} P(N_j \geq k + 1|X_0 = i) = \sum_{k=0}^{\infty} f_{ij} f_{jj}^k = \frac{f_{ij}}{1 - f_{jj}} \end{aligned}$$

The last statement follows from an application of (i):

$$P(N_j = k|X_0 = j) = (1 - f_{jj})f_{jj}^k$$

(iii) We compute this directly as

$$P(N_j = \infty|X_0 = j) = \lim_{k\to\infty} P(N_j \geq k) = \lim_{k\to\infty} f_{jj}^k = 1$$

□

## 2.2 Computation of $f_{ij}^{(n)}$

By definition, $f_{ij}^{(1)} = p_{ij}$ and

$$\begin{aligned}
f_{ij}^{(n)} &= P(X_1 \neq j, X_2 \neq j, ..., X_{n-1} \neq j, X_n \neq j | X_0 = i) \\
&= \sum_{k \in S, k \neq j} P(X_1 = k, X_2 \neq j, ..., X_{n-1} \neq j, X_n \neq j | X_0 = i) \\
&= \sum_{k \in S, k \neq j} P(X_2 \neq j, ..., X_{n-1} \neq j, X_n \neq j | X_0 = i, X_1 = k) P(X_1 = k | X_0 = i) \\
&= \sum_{k \in S, k \neq j} P(X_2 \neq j, ..., X_{n-1} \neq j, X_n \neq j | X_1 = k) P(X_1 = k | X_0 = i) \\
&= \sum_{k \in S, k \neq j} P(X_1 \neq j, ..., X_{n-1} \neq j, X_{n-1} \neq j | X_0 = k) P(X_1 = k | X_0 = i) \\
f_{ij}^{(n)} &= \sum_{k \in S, k \neq j} p_{ik} f_{kj}^{(n-1)}
\end{aligned}$$

*Remark* 2.5. Define the column vector $f^{(n)} = (f_{1j}^{(n)}, f_{2j}^{(n)}, ..., f_{ij}^{(n)}, ...f_{|S|j}^{(n)})^T$ and the matrix ${}^{(j)}P$ as the $P$ matrix with the $j^{th}$ column replaced by a column of zeroes. Then we can write

$$f^{(n)} = {}^{(j)}P f^{(n-1)} = {}^{(j)}P^{(n-1)} f^{(1)}$$

## 2.3 Periodicity

**Definition 2.9.** The **period** of a state $i$, $d(i)$, is defined as

$$d(i) = \gcd(n \geq 1 : p_{ii}^{(n)} > 0)$$

If $d(i) = 1$ then we say state $i$ is **periodic**. If $d(i) > 1$ then we say state $i$ has period $d(i)$.

**Example 2.10.** Let $\{X_k\}$ be a sequence of iid r.v.s with

$$P(X_k = 1) = p, P(X_k = -1) = q$$

with $p + q = 1$ and $0 < p, q < 1$. Define $S_0 = 0$ and $S_n = S_0 + \sum_{k=1}^{n} X_k$. Then $\{S_n : n \geq 0\}$ is a Markov chain with $S = \{..., -1, 0, 1, ...\}$and

$$P_{ij} = \begin{cases} q & i - j = 1 \\ p & i - j = -1 \\ 0 & \text{otherwise} \end{cases}$$

It is clear that $d(0) = 2$ since $p_{00}^{(n)} > 0$ for $n$ divisible by 2.

**Example 2.11.** Let $\{X_k\}$ be a sequence of iid r.v.s with

$$P(X_k = 1) = p, P(X_k = 0) = r, P(X_k = -1) = q$$

with $p + r + q = 1$ and $0 < p, r, q < 1$. Define $S_0 = 0$ and $S_n = S_0 + \sum_{k=1}^{n} X_k$. Then $d(0) = 1$ and state $0$ is aperiodic.

**Example 2.12.** Consider a Markov chain with $S = \{1, 2, 3\}$ and

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}$$

Since $p_{11}^{(2)}, p_{11}^{(3)} > 0$ then $d(1) = 1$.

## 2.4 Solidarity Properties

**Definition 2.10.** A property is called a **solidarity or equivalence property** if whenever state $i$ has a property and $i \leftrightarrow j$ then $j$ also has the same property. So if $C$ is an equivalence class and if $i \in C$ has a property, then all $j \in C$ has the same property.

**Proposition 2.8.** *Recurrence[1], transience[2], and periodicity[3] are equivalence class properties.*

*Proof.* [1,2] Suppose that $i \leftrightarrow j$ and $i$ is recurrent. Then, there exists $n$ such that $p_{ij}^{(n)} > 0$ and similarly there exists $m$ such that $p_{ji}^{(m)} > 0$. In order to prove that $j$ is recurrent, we will show $\sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty$. Then,

$$\begin{aligned} p_{jj}^{(n+k+m)} &= \sum_{\beta \in S} \sum_{\alpha \in S} p_{j\alpha}^{(m)} p_{\alpha\beta}^{(k)} p_{\beta j}^{(n)} \\ &\geq p_{ji}^{(m)} p_{ii}^{(k)} p_{ij}^{(n)} \\ &= c p_{ii}^{(k)}, c = p_{ji}^{(m)} p_{ij}^{(n)} > 0 \end{aligned}$$

Since $i$ is recurrent, then $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$ and hence

$$\sum_{l=0}^{\infty} p_{jj}^{(l)} \geq \sum_{k=0}^{\infty} p_{jj}^{(n+k+m)} \geq c \sum_{k=0}^{\infty} p_{ii}^{(k)} = \infty$$

The contrapositive tells us that transience is an equivalence property.

[3] Suppose $i \leftrightarrow j$ and $i$ has period $d(i)$ and $j$ has period $d(j)$ and from our previous result, we know $p_{jj}^{(n+k+m)} \geq c p_{ii}^{(k)}$. If $k = 0$ then $p_{ii}^{(k)} = 1$ and $p_{ii}^{(n+m)} \geq c > 0$ so $(n+m) = k_1 d(j)$. On the other hand, if $k$ is such that $p_{ii}^{(k)} > 0$ we have $p_{jj}^{(n+m+k)} \geq c p_{ii}^{(k)} > 0$. Then $(n+m+k) = k_2 d(j)$. Now,

$$k = (n+m+k) - (n+m) = (k_2 - k_1) d(j)$$

and so $d(j)$ is also a divisor of $\{n \geq 1 : p_{ii}^{(n)} > 0\}$. Then, $d(i) \geq d(j)$. Similarly, we can obtain $d(j) \geq d(i)$ since $\leftrightarrow$ is a symmetric relationship and hence $d(i) = d(j)$. □

**Example 2.13.** Going back to a recent example, let $\{X_k\}$ be a sequence of iid r.v.s with

$$P(X_k = 1) = p, P(X_k = 0) = r, P(X_k = -1) = q$$

with $p + r + q = 1$ and $0 < p, r, q < 1$. Define $S_0 = 0$ and $S_n = S_0 + \sum_{k=1}^{n} X_k$. Let us check that $\sum p_{00}^{(n)} = \infty$. Now since $p_{00}^{(2n+1)} = 0$ for $n \in \mathbb{N}$ and

$$\begin{aligned} p_{00}^{(2n)} = \binom{2n}{n} p^n (1-p)^n &= \frac{(2n)!}{n!n!} p^n (1-p)^n \\ &\approx \frac{\sqrt{2\pi} e^{-2n} (2n)^{2n+\frac{1}{2}} p^n (1-p)^n}{2\pi e^{-2n} n^{2n+1}} \\ &= \frac{(4p(1-p))^n}{\sqrt{\pi n}} \end{aligned}$$

using Stirling's approximation which states $n! \approx \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}$. Now for $p = \frac{1}{2}$ we have $p_{00}^{(2n)} \approx \frac{1}{\sqrt{\pi n}}$ which in the tail of the series defines a series larger than the Harmonic series and hence $\sum_{n \in \mathbb{N}} p_{00}^{(n)} = \infty$. We may repeat the same procedure for $p < \frac{1}{2}, p > \frac{1}{2}$ to see in these cases that $\sum_{n \in \mathbb{N}} p_{00}^{(n)} < \infty$. Hence, state 0 is transient and all states are transient. (For completeness, we can also repeat the above using the upper bound of Stirling's formula)

**Example 2.14.** Consider the Simple Branching process with $S = \{0, 1, 2, ...\}$, $P(Z_{ij} = k) = p_k$, $p_1 \neq 1$, and note that 0 is an absorbing state and hence it is recurrent. Assume $p_0 = 0$. Then

$$f_{kk} = P(Z_{n+1} = k | Z_n = k) = (p_1)^k < 1$$

and in this case, all states are transient. Suppose

$$p_0 = 1 \implies p_{k0} = 1 \implies f_{kk} = 0 \implies k \text{ is transient}$$

and hence all states are transient again. Now suppose that $0 < p_0 < 1$. Then,

$$\begin{aligned} f_{kk} \leq P(Z_1 \neq 0 | Z_0 = k) &= 1 - P(Z_1 = 0 | Z_0 = k) \\ &= 1 - (p_0)^k < 1 \end{aligned}$$

and so all states except 0 are transient in any type of branching process.

## 2.5 More State Space Decomposition

We can decompose the state space $S$ into $S = T \cup (\bigcup_i C_i)$ where $C_i's$ are closed sets of recurrent states, $T$ is a set of transient states (not necessarily in the same equivalence class).

**Proposition 2.9.** *Suppose $j$ is recurrent and for $k \neq j$ we have $j \to k$. Then,*

*(i) $k$ is recurrent*

*(ii) $j \leftrightarrow k$*

*(iii) $f_{jk} = f_{kj} = 1$*

*Proof.* (i) was proven in a previous lecture.

We first show (ii). This, we need to prove that $k \to j$. Suppose that $j$ is not accessible from $k$; that is

$$P(X_n \neq j, \forall n \geq 1 | X_0 = k) = 1$$

Since $j \to k$ there exists $m$ such that $p_{jk}^{(m)} > 0$ and since $j$ is recurrent, we also have $\sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty$. Now,

$$\begin{aligned} 0 &= P(X_l \neq j, \forall l \geq m | X_0 = j) \\ &\geq P(X_l \neq j, X_m = k, \forall l \geq m | X_0 = j) \\ &= P(X_m = k | X_0 = j) P(X_l \neq j, \forall l \geq m | X_0 = j) \\ &= p_{jk}^{(m)} \underbrace{P(X_l \neq j, l \geq 1 | X_0 = k)}_{=1} \\ &> 0 \end{aligned}$$

Thus, this is a contradiction and $j$ is accessible from $k$.

(iii) Since $j \leftrightarrow k$, there exists $m$ such that

$$P(X_1 \neq j, X_2 \neq j, ..., X_{m-1} \neq j, X_m = k | X_0 = j) > 0$$

Since $j$ is recurrent, we have $f_{jj} = 1$. Therefore,

$$\begin{aligned} 0 = 1 - f_{jj} &= P(\tau_j(1) = \infty | X_0 = j) \\ &\geq P(\tau_j(1) = \infty, X_m = k | X_0 = j) \\ &\geq P(X_1 \neq j, X_2 \neq j, ..., X_m = k, \tau_j(1) = \infty | X_0 = j) \\ &= P(\tau_j(1) = \infty | X_1 \neq j, X_2 \neq j, ..., X_{m-1} \neq j, X_0 = j, X_m = k) \times \\ & P(X_1 \neq j, X_2 \neq j, ..., X_{m-1} \neq j, X_m = k | X_0 = j) \\ &= \underbrace{P(\tau_j(1) = \infty | X_m = k)}_{1 - f_{kj}} \underbrace{P(X_1 \neq j, X_2 \neq j, ..., X_{m-1} \neq j, X_m = k | X_0 = j)}_{>0} \end{aligned}$$

and hence $1 - f_{kj} \leq 0 \implies f_{kj} = 1$. By symmetry, $f_{jk} = 1$ as well. $\square$

**Corollary 2.4.** *The state space $S$ of a Markov chain can be decomposed as*

$$S = T \cup C_1 \cup C_2 \cup ...$$

*where $T$ consists of transient states (not necessarily in one class) and $C_1, C_2, ...$ are closed disjoint classes of recurrent states. If $j \in C_\alpha$ then*

$$f_{jk} = \begin{cases} 1 & k \in C_\alpha \\ 0 & \text{otherwise} \end{cases}$$

*Furthermore, if we relabel the states so that for $i = 1, 2, ...$ states in $C_i$ have consecutive labels with states in $C_1$ having the smallest labels, $C_2$ the next smallest, etc. We can represent this as*

$$\begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ \vdots \\ T \end{array} \begin{array}{c} \begin{array}{ccccc} C_1 & C_2 & C_3 & \cdots & T \end{array} \\ \begin{pmatrix} P_1 & 0 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 & 0 \\ 0 & 0 & P_3 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Q_1 & Q_2 & Q_3 & \cdots & Q_T \end{pmatrix} \end{array}$$

*where $P_1, P_2, P_3$ are square stochastic matrices.*

*Remark* 2.6. If $S$ contains an infinite number of states, it is possible for $S = T$ as have seen in the simple random walk. If $S$ is finite however, not all states can be transient.

**Proposition 2.10.** *If $S$ is finite, not all states can be transient.*

*Proof.* Suppose that $S = \{0, 1, 2, ..., m\}$ and $S = T$. Let $j \in T$ and note that

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$$

for any $i \in S$. Now since $\sum_{j \in S} p_{ij}^{(n)}$ is the row sum of $P^{(n)}$ it is 1 and

$$1 = \lim_{n\to\infty} \sum_{j\in S} p_{ij}^{(n)} = \sum_{j\in S} \lim_{n\to\infty} p_{ij}^{(n)} = \sum_{j \in S} 0 = 0$$

which is impossible. □

**Example 2.15.** Consider $S = \{0, 1, 2, 4\}$ with

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Here, $C_1 = \{0\}, C_2\{4\}$ and $T = \{1, 2, 3\}$ since we may write

$$P = \begin{array}{c} \\ 0 \\ 4 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 4 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \end{pmatrix} \end{array}$$

**Example 2.16.** Consider $S = \{1, 2, 3, 4, 5\}$ with

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 6 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix} \end{array}$$

Drawing the probability transition diagram, we can see $C = \{1, 3, 5\}$ with $T = \{2, 4\}$.

## 2.6 Absorption Probabilities

**Definition 2.11.** Suppose that $S = T \cup C_1 \cup C_2 \cup \ldots$ and define $\tau = \inf\{n \geq 0 : X_n \notin T\}$ as the exit time from $T$. Of course it is possible that $P(\tau = \infty | X_0 = i) > 0$. Assume $P(\tau < \infty | X_0 = i) = 1$ and let

$$P = \begin{pmatrix} Q & R \\ 0 & P_2 \end{pmatrix}, Q = (Q_{ij}, i, j \in T), R = (R_{kl}, k \in T, l \in T^c)$$

When $\tau$ is finite, $X_\tau$ is the first state that the chain visits outside the transient states. Define

$$\begin{aligned} u_{ik} &= P(X_\tau = k | X_0 = i) \\ u_i(C_l) &= P(X_\tau \in C_l | X_0 = i) = \sum_{k \in C_l} u_{ik} \end{aligned}$$

*Remark* 2.7. We claim that $Q_{ij}^{(n)} = p_{ij}^{(n)}$. To see this, remark that

$$\begin{aligned} Q_{ij}^{(n)} &= \sum_{j_1, \ldots, j_{n-1} \in T} p_{ij_1} p_{j_1 j_2} \ldots p_{j_{n-1} j} \\ &= P(X_n = j, \tau > n | X_0 = i) \\ &= P(X_n = j | X_0 = i) \\ &= p_{ij}^{(n)} \end{aligned}$$

since $\{X_n = j\} \subset \{\tau > n\}$. From this, we can also see that $\sum_{n=0}^{\infty} Q_{ij}^{(n)} < \infty$. Now,

$$\begin{aligned}
u_{ij} &= P(X_\tau = j | X_0 = i) \\
&= \sum_{k \in S} P(X_\tau = j, X_1 = k | X_0 = i) \\
&= \sum_{k \in T} P(X_\tau = j, X_1 = k | X_0 = i) + \sum_{k \in T^c} P(X_\tau = j, X_1 = k | X_0 = i) \\
&= \sum_{k \in T} P(X_\tau = j, X_1 = k | X_0 = i) + p_{ij} \\
&= \sum_{k \in T} \sum_{n=2}^{\infty} P(\tau = n, X_\tau = j, X_1 = k | X_0 = i) + p_{ij} \\
&= \sum_{k \in T} \sum_{n=2}^{\infty} P(X_2 \in T, X_3 \in T, ..., X_{n-1} \in T, X_n = j, X_1 = k | X_0 = i) + p_{ij} \\
&= \sum_{k \in T} \sum_{n=2}^{\infty} P(X_2 \in T, X_3 \in T, ..., X_{n-1} \in T, X_n = j | X_1 = k) P(X_1 = k | X_0 = i) + p_{ij} \\
&= \sum_{k \in T} \sum_{n=2}^{\infty} P(X_2 \in T, X_3 \in T, ..., X_{n-1} \in T, X_n = j | X_1 = k) p_{ik} + p_{ij} \\
&= \sum_{k \in T} \sum_{n=2}^{\infty} P(\tau = n-1, X_{n-1} = j | X_1 = k) p_{ik} + p_{ij} \\
&= \sum_{k \in T} P(X_\tau = j | X_0 = k) p_{ik} + p_{ij} \\
u_{ij} &= \sum_{k \in T} p_{ik} u_{kj} + p_{ij}
\end{aligned}$$

Hence if $U = (u_{ij}, i \in T, j \in T^c)$ then $U = QU + R \implies U(I - Q) = R$ and if $(I - Q)^{-1}$ exists then

$$U = (I - Q)^{-1} R, (I - Q)^{-1} = \sum_{n=0}^{\infty} Q^n$$

This also implies that

$$(I - Q)_{ij}^{-1} = E\left[\sum_{n=0}^{\infty} 1(X_n = j | X_0 = i)\right] = \sum_{n=0}^{\infty} p_{ij}^{(n)}$$
$$= \text{expected \# of visits to } j \text{ from } i$$

# 3 Stationary Distributions

**Definition 3.1.** A stochastic process $\{Y_n : n \geq 0\}$ is **stationary** if of integers $m \geq 0$ and $k > 0$ we have

$$(Y_0, Y_1, ..., Y_m) \stackrel{d}{=} (Y_k, Y_{k+1}, ..., Y_{m+k})$$

Let $\pi = \{\pi_j : j \in S\}$ be a probability distribution. It is called a **stationary distribution** for the Markov chain with transition matrix $P$ if

$$\pi^T = \pi^T P, \pi_j = \sum_{k \in S} \pi_k P_{kj}, \forall j \in S$$

Let $P_\pi$ be the distribution of the chain when the initial distribution is $\pi$. That is,

$$P_\pi([\cdot]) = \sum_{i \in S} P([\cdot] | X_0 = i) \pi_i$$

**Proposition 3.1.** *With respect to $P_\pi$ we have that $\{X_n : n \geq 0\}$ is a stationary process. Thus,*

$$P_\pi(X_n = i_0, X_{n+1} = i_1, ..., X_{n+k} = i_k) = P_\pi(X_0 = i_0, X_1 = i_1, ..., X_k = i_k)$$

*for any $n \geq 0$, $k \geq 0$, and $i_0, i_1, ..., i_k \in S$. In particular, $P_\pi(X_n = j) = \pi_j$ for all $n \geq 0, j \in S$.*

*Proof.* We can compute directly

$$\begin{aligned}
&P_\pi(X_n = i_0, X_{n+1} = i_1, ..., X_{n+k} = i_k) \\
&= \sum_{i \in S} P(X_n = i_0, X_{n+1} = i_1, ..., X_{n+k} = i_k | X_0 = i) P(X_0 = i) \\
&= \sum_{i \in S} \pi_i p_{i i_0}^{(n)} p_{i_0 i_1} p_{i_1 i_2} ... p_{i_{k-1} i_k} \\
&= \pi_{i_0} p_{i_0 i_1} p_{i_1 i_2} ... p_{i_{k-1} i_k} \\
&= P_\pi(X_0 = i_0, X_1 = i_1, ..., X_k = i_k)
\end{aligned}$$

□

**Definition 3.2.** We call $\nu = \{\nu_j : j \in S\}$ an invariant measure if $\nu^T = \nu^T P$. If $\nu$ is an invariant measure and a probability distribution then it is is a stationary distribution.

**Proposition 3.2.** *Let $i \in S$ be recurrent and define for $j \in S$*

$$\nu_j = E\left[\sum_{0 \leq n \leq \tau_i(1)-1} 1(X_n = j) | X_0 = i\right] = \sum_{n=0}^{\infty} P(X_n = j, \tau_i(1) > n | X_0 = i)$$

*Then $\nu$ is an invariant measure. If $i$ is positive recurrent, then*

$$\pi_j = \frac{\nu_j}{E\left[\tau_i(1) | X_0 = i\right]}$$

*is a stationary distribution.*

*Proof.* We will first show that $\nu^T = \nu^T P$. Clearly $\nu_i = 1$. Now consider $j \neq i$. We need to show that $\nu_j = \sum_{k \in S} \nu_k p_{kj}$. Now since $X_{\tau_i(1)} = i$ and $X_0 = i$, then we have

$$\begin{aligned}
\nu_j &= E\left[\sum_{1\le n\le \tau_i(1)} 1(X_n = j)|X_0 = i\right] \\
&= E\left[\sum_{n=1}^{\infty} 1(X_n = j, \tau_i(1) \ge n)|X_0 = i\right] \\
&= \sum_{n=1}^{\infty} E\left[1(X_n = j, \tau_i(1) \ge n)|X_0 = i\right] \\
&= \sum_{n=1}^{\infty} P\left(X_n = j, \tau_i(1) \ge n|X_0 = i\right) \\
&= p_{ij} + \sum_{n=2}^{\infty} P\left(X_n = j, \tau_i(1) \ge n|X_0 = i\right) \\
&= p_{ij} + \sum_{n=2}^{\infty} \sum_{\substack{k \in S \\ k \ne i}} P\left(X_n = j, X_{n-1} = k, \tau_i(1) \ge n|X_0 = i\right) \\
&= p_{ij} + \sum_{n=2}^{\infty} \sum_{\substack{k \in S \\ k \ne i}} P\left(X_n = j|X_{n-1} = k, \tau_i(1) \ge n, X_0 = i\right) P(X_{n-1} = k, \tau_i(1) \ge n, X_0 = i) \\
&= p_{ij} + \sum_{n=2}^{\infty} \sum_{\substack{k \in S \\ k \ne i}} P\left(X_n = j|X_{n-1} = k, \tau_i(1) \ge n, X_0 = i\right) P(X_{n-1} = k, \tau_i(1) \ge n|X_0 = i) \\
&= p_{ij} + \sum_{n=2}^{\infty} \sum_{\substack{k \in S \\ k \ne i}} p_{kj} P(X_{n-1} = k, \tau_i(1) \ge n|X_0 = i)
\end{aligned}$$

Next, we observe that

$$\{\tau_i(1) \ge n, X_{n-1} = k\} = \{X_1 \ne i, X_2 \ne i, ..., X_{n-1} \ne i, X_{n-1} = k\}$$

and we may continue as

$$\begin{aligned}
\nu_j &= p_{ij} + \sum_{n=2}^{\infty} \sum_{\substack{k \in S \\ k \neq i}} p_{kj} P(X_{n-1} = k, \tau_i(1) \geq n | X_0 = i) \\
&= p_{ij}\nu_i + \sum_{n=2}^{\infty} \sum_{\substack{k \in S \\ k \neq i}} p_{kj} P(X_{n-1} = k, \tau_i(1) \geq n | X_0 = i) \\
&= p_{ij}\nu_i + \sum_{\substack{k \in S \\ k \neq i}} \sum_{n=2}^{\infty} p_{kj} P(\tau_i(1) \geq n, X_{n-1} = k | X_0 = i) \\
&= p_{ij}\nu_i + \sum_{\substack{k \in S \\ k \neq i}} \sum_{n=1}^{\infty} p_{kj} P(\tau_i(1) \geq n+1, X_n = k | X_0 = i) \\
&= p_{ij}\nu_i + \sum_{\substack{k \in S \\ k \neq i}} p_{kj} \sum_{n=1}^{\infty} P(\tau_i(1) \geq n+1, X_n = k | X_0 = i) \\
&= p_{ij}\nu_i + \sum_{\substack{k \in S \\ k \neq i}} p_{kj} \sum_{n=1}^{\infty} E\left[1(\tau_i(1) \geq n+1, X_n = k) | X_0 = i\right] \\
&= p_{ij}\nu_i + \sum_{\substack{k \in S \\ k \neq i}} p_{kj} E\left[\sum_{0 \leq n \leq \tau_i(1)-1} 1(X_n = k) | X_0 = i\right] \\
&= p_{ij}\nu_i + \sum_{\substack{k \in S \\ k \neq i}} p_{kj}\nu_k = \sum_{k \in S} p_{kj}\nu_k
\end{aligned}$$

So $\nu_j$ is an invariant measure. Next, we calculate

$$\begin{aligned}
\sum_{j \in S} \nu_j &= \sum_{j \in S} E\left[\sum_{0 \leq n \leq \tau_i(1)-1} 1(X_n = j) | X_0 = i\right] \\
&= E\left[\sum_{j \in S} \sum_{0 \leq n \leq \tau_i(1)-1} 1(X_n = j) | X_0 = i\right] \\
&= E\left[\sum_{0 \leq n \leq \tau_i(1)-1} \sum_{j \in S} 1(X_n = j) | X_0 = i\right] \\
&= E\left[\sum_{n=0}^{\tau_i(1)-1} 1 | X_0 = i\right] = E\left[\tau_i(1) | X_0 = i\right]
\end{aligned}$$

and we are done as the normalized $\nu_j$ is $\pi_j$. □

**Proposition 3.3.** *If the Markov chain is irreducible and recurrent, then an invariant measure $\nu$ exists and satisfies $0 < \nu_j < \infty, \forall j \in S$ and $\nu$ is unique up to a constant. If $\nu_1^T = \nu_1^T P$ and $\nu_2^T = \nu_2^T P$ then $\nu_1 = c\nu_2$. Furthermore, if the Markov chain is* *positive recurrent and irreducible, there exists a unique stationary distribution* $\pi$ *where*

$$\pi_j = \frac{1}{E[\tau_j(1)|X_0 = j]}$$

**Lemma 3.1.** *(Strong Law of Large Numbers (SLLN)) Suppose* $\{Y_n\}$ *is a sequence of iid r.vs with* $E(|Y_i|) < \infty$*. Then,*

$$P\left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n} Y_i}{n} = E[Y_1]\right) = 1$$

*(converges almost surely (a.s.)).*

**Proposition 3.4.** *Suppose the Markov chain is irreducible and positive recurrent, and let* $\pi$ *be the unique stationary distribution. Then*

$$\lim_{N\to\infty} \frac{\sum_{n=0}^{N} f(X_n)}{N} = \sum_{j\in S} f(j)\pi_j, \text{ a.s. } \implies P\left(\lim_{N\to\infty} \frac{\sum_{n=0}^{N} f(X_n)}{N} = \sum_{j\in S} f(j)\pi_j\right) = 1$$

*Note that if* $f(k) = 1(k = i)$ *then*

$$\lim_{N\to\infty} \frac{\sum_{n=0}^{N} f(X_n)}{N} = \pi_i$$

*Proof.* Remark that if $f$ is non-negative ($f \geq 0$), then

$$\begin{aligned}
\sum_{j\in S} f(j)\pi_j &= \sum_{j\in S} f(j) \frac{E\left[\sum_{n=0}^{\tau_i(1)-1} 1(X_n = j)|X_0 = i\right]}{E\left[\tau_i(1)|X_0 = i\right]} \\
&= \frac{E\left[\sum_{j\in S}\sum_{n=0}^{\tau_i(1)-1} f(j)1(X_n = j)|X_0 = i\right]}{E\left[\tau_i(1)|X_0 = i\right]} \\
&= \frac{E\left[\sum_{n=0}^{\tau_i(1)-1}\sum_{j\in S} f(j)1(X_n = j)|X_0 = i\right]}{E\left[\tau_i(1)|X_0 = i\right]} \\
&= \frac{E\left[\sum_{n=0}^{\tau_i(1)-1} f(X_n)|X_0 = i\right]}{E\left[\tau_i(1)|X_0 = i\right]} \overset{(*)}{=} \frac{E\left[\sum_{n=1}^{\tau_i(1)} f(X_n)|X_0 = i\right]}{E\left[\tau_i(1)|X_0 = i\right]}
\end{aligned}$$

where (*) is because $X_0 = X_{\tau_i(1)} = i$. Now define $B(N) = \sup\{k \geq 0 : \tau_i(k) \leq N\}$, the number of visits to $i$ before time $N$, and $\eta_k = \sum_{n=\tau_i(k)+1}^{\tau_i(k+1)} f(X_n)$. The sequence $\{\eta_k\}$ is a sequence of iid r.vs (times between Markov processes starting from the same state are independent) and

$$\lim_{m\to\infty} \frac{1}{m}\sum_{k=1}^{m} \eta_k = E\left[\sum_{n=1}^{\tau_i(1)} f(X_n)|X_0 = i\right]$$

Next, remark that

$$\sum_{n=0}^{\tau_i(B(N))} f(X_n) \leq \sum_{n=0}^{N} f(X_n) \leq \sum_{n=0}^{\tau_i(B(N)+1)} f(X_n)$$

with lower bound

$$\begin{aligned}
\sum_{n=0}^{\tau_i(B(N))} f(X_n) &= \sum_{n=0}^{\tau_i(1)} f(X_n) + \sum_{n=\tau_i(1)+1}^{\tau_i(B(N))} f(X_n) \\
&= \sum_{n=0}^{\tau_i(1)} f(X_n) + \sum_{k=1}^{B(N)-1} \eta_k
\end{aligned}$$

and similarly upper bound of

$$\sum_{n=0}^{\tau_i(B(N))+1} f(X_n) = \sum_{n=0}^{\tau_i(1)} f(X_n) + \sum_{k=1}^{B(N)} \eta_k$$

Looking at the limiting behaviour:

$$\lim_{N\to\infty}\left(\underbrace{\frac{\sum_{n=0}^{\tau_i(1)} f(X_n)}{N}}_{\to 0}+\frac{\sum_{k=1}^{B(N)-1}\eta_k}{N}\right)\le \lim_{N\to\infty}\frac{\sum_{n=0}^{N} f(X_n)}{N}\le \lim_{N\to\infty}\left(\underbrace{\frac{\sum_{n=0}^{\tau_i(1)} f(X_n)}{N}}_{\to 0}+\frac{\sum_{k=1}^{B(N)}\eta_k}{N}\right)$$

Now,

$$\lim_{N\to\infty}\frac{\sum_{k=1}^{B(N)}\eta_k}{N}=\lim_{N\to\infty}\frac{\sum_{k=1}^{B(N)}\eta_k}{N}\cdot\frac{B(N)}{B(N)}\overset{(?)}{=}\frac{E[\eta_1|X_0=i]}{E[\tau_i(1)|X_0=i]}$$

and similarly

$$\lim_{N\to\infty}\frac{\sum_{k=1}^{B(N)-1}\eta_k}{N}=\lim_{N\to\infty}\frac{\sum_{k=1}^{B(N)}\eta_k}{N}\cdot\frac{B(N)-1}{B(N)-1}\overset{(?)}{=}\frac{E[\eta_1|X_0=i]}{E[\tau_i(1)|X_0=i]}$$

where $(?)$ comes from the fact that

$$\lim_{N\to\infty}\frac{\sum_{k=1}^{B(N)}\eta_k}{B(N)}=E[\eta_1|X_0=i]$$

from the SLLN and

$$\lim_{N\to\infty}\frac{B(N)}{N}=\frac{1}{E[\tau_i(i)|X_0=i]}$$

comes from the fact that

$$\begin{aligned}
\tau_i(B(N))\le N\le \tau_i(B(N)+1) &\implies \frac{\tau_i(B(N))}{B(N)}\le\frac{N}{B(N)}\le\frac{\tau_i(B(N)+1)}{B(N)}\cdot\frac{B(N)+1}{B(N)+1}\\
&\implies \lim_{N\to\infty}\frac{\tau_i(B(N))}{B(N)}\le\frac{N}{B(N)}\le\lim_{N\to\infty}\frac{\tau_i(B(N)+1)}{B(N)}\cdot\frac{B(N)+1}{B(N)+1}\\
&\implies E[\tau_i(1)|X_0=i]\le\frac{N}{B(N)}\le E[\tau_i(1)|X_0=i]\cdot 1
\end{aligned}$$

Hence we finally have

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N} f(X_n)}{N}=\frac{E\left[\sum_{n=1}^{\tau_i(1)} f(X_n)|X_0=i\right]}{E\left[\tau_i(1)|X_0=i\right]}=\sum_{j\in S} f(j)\pi_j$$

□

**Corollary 3.1.** *If $f$ is bounded then*

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N} E\left[f(X_n)|X_0=i\right]}{N}=\sum_{j\in S} f(j)\pi_j$$

*In particular, if $f(k)=1[k=j]$ then we have $E\left[f(X_n)|X_0=i\right]=P(X_n=j|X_0=i)=p_{ij}^{(n)}$. So,*

$$\lim_{N\to\infty}\frac{\sum_{n=1}^{N} p_{ij}^{(n)}}{N}=\pi_j\implies\lim_{N\to\infty}\frac{\sum_{n=1}^{N} P^n}{N}=\Pi$$

*Proof.* We know that

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N} f(X_n)}{N}=\sum_{j\in S} f(j)\pi_j,\ \text{a.s.}$$

and suppose that $|f(k)|\le M$ for all $k\in S$. That is,

$$\left|\frac{\sum_{n=0}^{N} f(X_n)}{N}\right|\le M$$

Then, by the dominated convergence theorem

$$
\begin{aligned}
E\left[\lim_{N\to\infty}\frac{\sum_{n=0}^{N} f(X_n)}{N}|X_n=0\right] &= \lim_{N\to\infty}\frac{E\left[\sum_{n=0}^{N} f(X_n)|X_n=0\right]}{N} \\
&= \lim_{N\to\infty}\sum_{n=0}^{N}\frac{E\left[f(X_n)|X_n=0\right]}{N} \\
&= \sum_{j\in S} f(j)\pi_j
\end{aligned}
$$

□

## 3.1 Limiting Distribution

**Proposition 3.5.** *A limit distribution is a stationary distribution.*

*Proof.* Directly, we have

$$
\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty}\sum_{k\in S} p_{ik}^{(n)} p_{kj}
$$

Suppose that $S = \{0, 1, 2, ...\}$. Remark that for all $M \in \mathbb{N}$,

$$
\pi_j \geq \lim_{n\to\infty}\sum_{k=0}^{M} p_{ik}^{(n)} p_{kj} = \sum_{k=0}^{M}\pi_k p_{kj}
$$

Suppose there exists some $j'$ such that

$$
\pi_{j'} > \sum_{k=0}^{\infty}\pi_k p_{kj'} \implies \sum_{j=0}^{\infty}\pi_j > \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\pi_k p_{kj} = \sum_{k=0}^{\infty}\pi_k\sum_{j=0}^{\infty}p_{kj} = \sum_{k=0}^{\infty}\pi_k = 1
$$

which is impossible. Thus, we have

$$
\pi_j = \sum_{k=0}^{\infty} p_{kj}\pi_k
$$

□

**Theorem 3.1.** *Suppose the Markov chain is irreducible and aperiodic and that a stationary distribution $\pi$ exists with*

$$
\pi^T = \pi^T P \text{ and } \sum_{j\in S}\pi_j = 1 \text{ with } \pi_j \geq 0
$$

*Then:*

*(1) The Markov chain is positive recurrent*

*(2) $\pi$ is a limit distribution with $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \pi_j, \forall i, j \in S$*

*(3) For all $j \in S$, $\pi_j > 0$*

*(4) The stationary distribution is unique*

*Proof.* (1) If the chain were transient then

$$
\lim_{n\to\infty} p_{ij}^{(n)} = 0, \forall i, j \in S
$$

and so $\pi_j = \sum_{i\in S}\pi_i p_{ij}^{(n)} \to 0$ for all $j \in S$. But if $\pi_j = 0, \forall j \in S$ then we cannot have $\sum_{j\in S}\pi_j = 1$. Now,

$$
\nu_j = E\left[\sum_{1\leq n\leq\tau_i(1)} 1(X_n = j)|X_0 = i\right] = c\pi_j
$$

and thus

$$\begin{aligned}
\infty > \sum_{j\in S} \nu_j = \sum_{j \in S} E\left[\sum_{1\le n\le \tau_i(1)} 1(X_n = j)|X_0 = i\right] \\
&= E\left[\sum_{j\in S}\sum_{1\le n\le \tau_i(1)} 1(X_n = j)|X_0 = i\right] \\
&= E\left[\sum_{1\le n\le \tau_i(1)}\sum_{j\in S} 1(X_n = j)|X_0 = i\right] \\
&= E\left[\sum_{1\le n\le \tau_i(1)} 1|X_0 = i\right] \\
&= E\left[\tau_i(1)|X_0 = i\right]
\end{aligned}$$

So the chain is positive recurrent. This gives us

$$\pi_j = \frac{1}{E\left[\tau_j(1)|X_0 = j\right]}$$

and (3) and (4) follow from a previous proposition and the above remark.

The proof of (2) is much more involved. We first start with a lemma. □

**Lemma 3.2.** *Let the chain be irreducible and aperiodic. Then for* $i, j \in S$ *there exists* $n_0(i,j)$ *such that for all* $n \ge n_0(i,j)$ *we have* $p_{ij}^{(n)} > 0$*.*

*Proof.* Define $\Lambda = \{n : p_{jj}^{(n)} > 0\}$.

(1) We know that the greatest common divisor of the set $\Lambda$ is 1.

(2) If $m \in \Lambda, n \in \Lambda$ then $m + n \in \Lambda$ by the fact that

$$p_{jj}^{(n+m)} = \sum_{k\in S} p_{jk}^{(n)} p_{kj}^{(m)} \ge p_{jj}^{(n)} p_{jj}^{(m)} > 0$$

Then $\Lambda$ contains all sufficiently large integers, say $n \ge n_1$, such that $p_{jj}^{(n)} > 0$. So given $i, j \in S$ there exists $r$ such that $p_{ij}^{(r)} > 0$. In order to see this, for $n \ge r + n_1$ we have

$$p_{ij}^{(n)} = \sum_{k\in S} p_{ij}^{(r)} p_{kj}^{(n-r)} \ge p_{ij}^{(r)} p_{jj}^{(n-r)} > 0$$

by choice of $r$ and $n - r \ge n_1$. □

*Proof.* [using "**coupling**"] (of (2) in the previous theorem) Let $\{X_n\}$ be the original Markov chain, and $\{Y_n\}$ be independent of $\{X_n\}$ and the same transition matrix as $\{X_n\}$ but the initial distribution of $\{Y_n\}$ is $\pi$. So $P(Y_n = j) = \pi_j$ for any $n \in \mathbb{N}$. Define $\varepsilon_n = (X_n, Y_n)$ so that $\{\varepsilon_n\}$ is a Markov chain with states in $S \times S$. Now,

$$\begin{aligned}
P(\varepsilon_{n+1} = (k,l)|\varepsilon_n = (i,j)) &= p_{ik}p_{jl} \\
P(\varepsilon_{n+1} = (k,l)|\varepsilon_0 = (i,j)) &= p_{ik}^{(n)} p_{jl}^{(n)}
\end{aligned}$$

and there exists $n_1, n_2$ such that $\forall n \ge n_1$ and $\forall m \ge n_2$, $p_{ik}^{(n)} > 0$ and $p_{jk}^{(m)} > 0$. Then for all $n \ge \max(n_1, n_2)$, we have $p_{ij}^{(n)} p_{jl}^{(n)} > 0$. Thus, $\{\varepsilon_n\}$ is an irreducible Markov chain.

Define $\pi_{k,l} = \pi_k \pi_l$. Then the product of the stationary distributions is a stationary distribution for $\{\varepsilon_n\}$:

$$\begin{aligned}\sum_{(i,j)\in S\times S} \pi_{i,j} P(\varepsilon_{n+1} = (k,l)|\varepsilon_n = (i,j)) &= \sum_{(i,j)\in S\times S} \pi_i \pi_j p_{ik} p_{jl} \\ &= \sum_{i\in S} \pi_i p_{ik} \sum_{j\in S} \pi_j p_{jl} \\ &= \pi_k \pi_l = \pi_{k,l}\end{aligned}$$

and since $\sum_{l\in S}\sum_{k\in S}\pi_{k,l} = \sum_{k\in S}\pi_k \sum_{l\in S}\pi_l = 1$ then $\{\varepsilon_n\}$ is positive recurrent. Define for $i_0 \in S$,

$$\tau_{i_0,i_0} = \inf\{n \geq 0 : \varepsilon_n = (i_0, i_0)\}$$

with the fact that $P(\tau_{i_0,i_0} < \infty) = 1$ (from recurrence of $\{\varepsilon_n\}$. Now,

$$\begin{aligned}P(X_n = j, \tau_{i_0,i_0} \leq n) &= \sum_{m=0}^{n} P(X_n = j, \tau_{i_0,i_0} = n) \\ &= \sum_{k\in S}\sum_{m=0}^{n} P(\varepsilon_n = (j,k), \tau_{i_0,i_0} = m) \\ &= \sum_{k\in S}\sum_{m=0}^{n} P(\varepsilon_n = (j,k)|\tau_{i_0,i_0} = m)P(\tau_{i_0,i_0} = m) \\ &= \sum_{k\in S}\sum_{m=0}^{n} P(\varepsilon_n = (j,k)|\varepsilon_m = (i_0,i_0))P(\tau_{i_0,i_0} = m) \\ &= \sum_{k\in S}\sum_{m=0}^{n} P(\varepsilon_{n-m} = (j,k)|\varepsilon_0 = (i_0,i_0))P(\tau_{i_0,i_0} = m) \\ &= \sum_{k\in S}\sum_{m=0}^{n} p_{i_0,j}^{(n-m)} p_{i_0,k}^{(n-m)} P(\tau_{i_0,i_0} = m) \\ &= \sum_{m=0}^{n} p_{i_0,j}^{(n-m)} P(\tau_{i_0,i_0} = m) \underbrace{\sum_{k\in S} p_{i_0,k}^{(n-m)}}_{=1} \\ P(X_n = j, \tau_{i_0,i_0} \leq n) &= \sum_{m=0}^{n} p_{i_0,j}^{(n-m)} P(\tau_{i_0,i_0} = m)\end{aligned}$$

By a similar construction, we can also show that

$$P(Y_n = j, \tau_{i_0,i_0} \leq n) = \sum_{m=0}^{n} p_{i_0,j}^{(n-m)} P(\tau_{i_0,i_0} = m)$$

Next, if we suppose that $X_0 = i$, then

$$\begin{aligned}|p_{ij}^{(n)} - \pi_j| &= |P(X_n = j) - P(Y_n = j)| \\ &= |P(X_n = j, \tau_{i_0,i_0} \leq n) + P(X_n = j, \tau_{i_0,i_0} > n) \\ &\qquad - P(Y_n = j, \tau_{i_0,i_0} \leq n) - P(Y_n = j, \tau_{i_0,i_0} > n)| \\ &= |P(X_n = j, \tau_{i_0,i_0} > n) - P(Y_n = j, \tau_{i_0,i_0} > n)| \\ &= |E[1(X_n = j)1(\tau_{i_0,i_0} > n)] - E[1(Y_n = j)1(\tau_{i_0,i_0} > n)]| \\ &= |E\left[[1(X_n = j) - 1(Y_n = j)]\,1(\tau_{i_0,i_0} > n)\right]| \\ |p_{ij}^{(n)} - \pi_j| &\leq E[1(\tau_{i_0,i_0} > n)] = P(\tau_{i_0,i_0} > n)\end{aligned}$$

Taking limits on $n \to \infty$ for both sides yields:

$$\lim_{n\to\infty} |p_{ij}^{(n)} - \pi_j| = 0$$

□

**Definition 3.3.** An irreducible, aperiodic, positive recurrent Markov chain is called an **ergodic** Markov chain.

**Corollary 3.2.** *Assume that a Markov chain is irreducible and aperiodic. A stationary distribution exists if and only if the chain is positive recurrent if and only if a limit distribution (defined through* $\lim_{n\to\infty} P^n$*) exists.*

*If the chain is irreducible and periodic, existence of a stationary distribution is equivalent to positive recurrent states.*

**Proposition 3.6.** *If the Markov chain is irreducible and aperiodic and either null recurrent or transient, then*

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0, \textit{ for all } i,j \in S$$

*We can conclude that in a finite state irreducible Markov chain, no state can be null recurrent.*

**Example 3.1.** Consider the inventory example with $X_n = X(\tau_n^+)$ where $\tau_n^+$ is right after the $n^{th}$ departure . Define $X_{n+1} = \max(X_n - 1, 0) + A_{n+1}$ where $A_{n+1}$ is the number of arrivals during the $(n+1)^{th}$ service time and $\{A_n\}$ a sequence of iid r.vs. Denote $P(A_1 = k) = a_k$ for $k = 0, 1, 2, ...$ and note that, starting from state 0,

$$P = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & a_3\cdots \\ 0 & 0 & a_0 & a_1 & a_2 a_3 \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

This gives us the following sequence of equations for the stationary distribution:

$$\begin{aligned} \pi_0 &= a_0\pi_0 + a_0\pi_1 \\ \pi_1 &= a_1\pi_0 + a_1\pi_1 + a_0\pi_2 \\ &\vdots \\ \pi_n &= a_n\pi_0 + \sum_{j=1}^{n+1} a_{n+1-j}\pi_j \\ \sum_{n=0}^{\infty} \pi_n &= 1 \end{aligned}$$

Using the generating series $\Pi(s) = \sum_{n=0}^{\infty} s^n\pi_n$, $A(s) = \sum_{n=0}^{\infty} s^n a_n$ we have

$$\begin{aligned} \Pi(s) = \sum_{n=0}^{\infty} s^n\pi_n &= \pi_0 \sum s^n a_n + \sum_{n=0}^{\infty} s^n \sum_{j=1}^{n+1} a_{n+1-j}\pi_j \\ &= \pi_0 A(s) + \sum_{j=1}^{\infty} \pi_j s^{j-1} \underbrace{\sum_{n=j-1}^{\infty} a_{n+1-j} s^{n-j+1}}_{A(s)} \\ &= \pi_0 A(s) + \frac{1}{s}\sum_{j=1}^{\infty} \pi_j s^j A(s) \\ &= \pi_0 A(s) + \frac{1}{s}(\Pi(s) - \pi_0)A(s) \end{aligned}$$

and hence

$$\Pi(s) = \frac{\pi_0 A(s)\left(1 - \frac{1}{s}\right)}{\frac{s - A(s)}{s}} = \frac{\pi_0 A(s)}{\frac{A(s)-s}{1-s}}$$

Using the fact that $\Pi(1) = 1$, we evaluate $\lim_{s\to 1}\Pi(s)$. First, using l'Hopital's rule,

$$\lim_{s\to 1}\frac{1-A(s)}{1-s} = A'(1) = \sum_{k=0}^{\infty} k a_k = \rho$$

and so the limit becomes

$$\lim_{s\to 1}\Pi(s) = 1 = \frac{\pi_0}{1-\rho} \implies \pi_0 = 1-\rho, \rho < 1$$

with existence requiring that $\rho < 1$.

# 4 Renewal Theory

**Definition 4.1.** Suppose that $\{Y_n : n \geq 0\}$ is a sequence of independent non-negative random variables. Furthermore, suppose the sequence $\{Y_n : n \geq 1\}$ is iid with common distribution $F(\cdot)$. We assume for all $n \geq 1$

$$P(Y_n < 0) = 0 \text{ and } P(Y_n = 0) < 1$$

For $n \geq 0$, define $S_n = Y_0 + Y_1 + ... + Y_n$. The sequence $\{S_n : n \geq 0\}$ is called a **renewal process**. The process is called **delayed** if $P(Y_0 > 0) > 0$ and **pure** if $S_0 = Y_0 = 0$. If $F(\infty) = 1$ then the process is called a **proper renewal process**. If $F(\infty) < 1$ then the process is called **terminating** or **transient**.

**Example 4.1.**

1) Replacement times of a machine where the lifetimes are independent identically distributed random variables.

2) Suppose $\{X_n : n \geq 0\}$ is a Markov chain with finite state space $S$. Fix state $i$ and define

$$\tau_0(i) = \inf\{n \geq 0 : X_n = i\}$$
$$\tau_{n+1}(i) = \inf\{n \geq \tau_n(i) : X_n = i\}$$

Then $\{\tau_n(i) : n \geq 0\}$ is a renewal process. If $X_0 = i$, it is a pure renewal process. Otherwise, it is a delayed renewal process.

3) A machine is either up or down. The sequence of on times are iid r.vs and the sequence of off times are iid r.vs.

**Definition 4.2.** Define $N(t) = \sum_{n=0}^{\infty} 1_{[0,t]}(S_n)$. We call $\{N(t) : t \geq 0\}$ a **counting process** and $U(t) = E[N(t)]$ a **renewal function**.

**[Review your Lebesgue-Stieltjes integrals more here]**

(1) If $U(x)$ is **a**bsolutely **c**ontinuous, then

$$\int_0^{\infty} g(x)dU(x) = \int_0^{\infty} g(x)U(dx) = \int_0^{\infty} g(x)u(x)dx$$

for some density function $u(x)$ where $U(b) - U(a) = \int_a^b u(s)ds$.

2) Suppose that $U$ is **d**iscrete. Then $\lim_{h\to 0} U(a_i + h) - U(a_i - h) = U(a_i) = w_i$. Thus, $U$ has atoms at locations $\{a_i\}$ of weight $\{w_i\}$. Then

$$\int_0^{\infty} g(x)U(dx) = \int_0^{\infty} g(x)dU(x) = \sum_i g(a_i)w_i$$

3) Suppose we have a mixed measure $U(x) = \alpha U_{AC}(x) + \beta U_D(x)$ for $\alpha, \beta > 0$. Then

$$\int g(x)U(x) = \alpha \int g(x)u_{AC}(x) + \beta \sum g(a_i)w_i$$

*Remark* 4.1. Consider the case where $U_{AC}(x) = \int_0^x u(s)\ ds$ and $U_d(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$ where for $x > 0$ we have $U(x) =$

$U_{AC}(x) + U_d(x) = 1 + \int_0^x u(s)\,ds$. Then,

$$\int_0^\infty g(x)U(dx) = g(0) + \int_0^\infty g(x)u(x)dx$$

## 4.1 Convolution

Suppose all functions are defined on $[0, \infty)$. A function $g$ is called **locally bounded** if $g$ is bounded on finite intervals. For a locally bounded non-negative function $g$ and a non-negative distribution function $F$ define the **convolution** of $F$ and $g$ as

$$F * g(t) = \int_0^t g(t-x)F(dx), \text{ for } t \geq 0$$

Here are some properties:

1. $F * g(t) \geq 0$ for all $t \geq 0$

2. $F * g(t)$ is locally bounded because for $0 \leq s \leq t$:

$$\begin{aligned}|F * g(s)| &= \left|\int_0^s g(s-x)F(dx)\right| \\ &\leq \int_0^s |g(s-x)|\,F(dx) \\ &\leq \int_0^s \sup_{0\leq s\leq t} g(s-x)F(dx) \\ &= \sup_{0\leq s\leq t} |g(s)|F(s)\end{aligned}$$

and hence $\sup_{0\leq s\leq t} |F * g(s)| \leq \sup_{0\leq s\leq t} |g(s)|F(t)$.

3. If $g$ is bounded and continuous, then $F * g$ is bounded and continuous. To see this, suppose that $Y_1$ is the random variable with distribution $F$. Then

$$F * g(t) = \int_0^t g(t-x)F(dx) = E[g(t-Y_1)]$$

If $t_n \to t$ then $g(t_n - Y_1) \to g(t - Y_1)$ almost surely from the Central Limit Theorem (CLT). From dominated convergence, we have

$$E[g(t_n - Y_1)] \to E[g(t - Y_1)]$$

4. The convolution can be repeated $F * (F * g)$ where

$$\begin{aligned}F^{0*}(x) &= 1_{[0,\infty)}^{(x)} \\ F^{1*}(x) &= F(x) \\ F^{2*}(x) &= F * F(x) \\ &\vdots \\ F^{n*}(x) &= F * F * ...F(x)\end{aligned}$$

5. Let $X_1$ and $X_2$ be two independent random variables with distributions $F_1$ and $F_2$. Then $F_1 * F_2$ is the distribution of

$X_1 + X_2$. To see this, note that

$$P(X_1 + X_2 \le t) = \iint_{\{(x,y)\in\mathbb{R}_+^2 : x+y\le t\}} F_1(dx)F_2(dy)$$
$$= \int_0^t \int_0^{t-y} F_1(dx)F_2(dy)$$
$$= \int_0^t F_1(t-y)F_2(dy)$$

6. $F_1 * F_2(t) = F_2 * F_1(t)$

7. Suppose that $Y_1, Y_2, ..., Y_2$ are iid r.vs with distribution function $F$. Then $F^{n*}$ is the distribution of $Y_1 + Y_2 + ... + Y_n$.

8. Suppose that $F_1$and $F_2$ are absolutely continuous with density functions $f_1$ and $f_2$ respectively. Then $F_1 * F_2$ is absolutely continuous with density function

$$f_1 * f_2 = \int_0^t f_1(t-y)f_2(y)\ dy$$

To see this, note that

$$F_1 * F_2(t) = \iint_{\{(x,y)\in\mathbb{R}_+^2 : x+y\le t\}} f_1(x)dx f_2(y)dy$$
$$= \int_0^t \int_0^{t-y} f_1(x)dx\ f_2(y)dy$$
$$= \int_0^t \int_y^t f_1(u-y)du\ f_2(y)dy$$
$$= \int_0^t \int_0^u f_1(u-y)f_2(y)\ dy\ du$$
$$= \int_0^t f_1 * f_2(u)\ du$$

In fact if $F$ is absolutely continuous, then for any function $G$, $F * G$ is absolutely continuous. To see this, suppose that $F$ has density function $f_1$. Then,

$$F * G(t) = \int_0^t \int_0^u f_1(u-y)G(dy)\ dy$$
$$= \int_0^t f_1 * G(y)\ dy$$

## 4.2 Laplace Transform

Suppose $X$ is a non-negative random variable with distribution function $F$. The **Laplace (Laplace-Stieltjes) transform** of $X$ or $F$ is

$$\hat{F}(\lambda) = E[e^{-\lambda X}] = \int_0^\infty e^{-\lambda x}F(dx), \lambda \ge 0$$

1. The Laplace transform uniquely determines the distribution function.

2. Suppose that $X_1$ and $X_2$ are iid r.vs with distribution functions $F_1$ and $F_2$ respectively. Then,

$$(\widehat{F_1 * F_2})(\lambda) = E[e^{-\lambda(X_1+X_2)}] = E[e^{-\lambda X_1}]E[e^{-\lambda X_2}] = \hat{F}_1(\lambda)\hat{F}_2(\lambda)$$

In general, $(\widehat{F^{n*}})(\lambda) = (\hat{F}(\lambda))^n$.

3. We have

$$(-1)^n \frac{d^n \hat{F}(\lambda)}{d\lambda^n} = \int_0^\infty e^{-\lambda x}x^n F(dx) \implies \lim_{\lambda\to 0}(-1)^n \frac{d^n \hat{F}(\lambda)}{d\lambda^n} = \int_0^\infty x^n F(dx)$$

and hence $E[X] = -\hat{F}(\lambda), E[X^2] = \hat{F}''(0)$.

4. We have

$$\int_0^\infty e^{-\lambda x} F(x) dx = \frac{1}{\lambda}\hat{F}(\lambda)$$

from the fact that

$$\begin{aligned}\int_0^\infty e^{-\lambda x} F(x) dx &= \int_0^\infty e^{-\lambda x} \int_0^x F(du)\ dx \\ &= \int_0^\infty F(du) \int_u^\infty e^{-\lambda x}\ dx \\ &= \int_0^\infty \frac{1}{\lambda} e^{-\lambda u} F(du) \\ &= \frac{1}{\lambda}\hat{F}(\lambda)\end{aligned}$$

and so $\int_0^\infty (1 - F(x))e^{-\lambda x} = \frac{1}{\lambda}(1 - \hat{F}(\lambda))$.

*Remark* 4.2. The Laplace transform can be defined for a general non-decreasing function $U$ on $[0, \infty)$ if there exist $a$ such that

$$\int_0^\infty e^{-\lambda x} U(dx) < \infty, \lambda > a$$

Then we say $\hat{U}(\lambda) = \int_0^\infty e^{-\lambda x} U(dx)$ for $\lambda > a$.

## 4.3 Renewal Functions

*Remark* 4.3. If $N(t) = \sum_{n=0}^\infty 1_{[0,t]}(S_n)$ and $E[N(t)] = U(t)$, then if $S_0 = 0$ we have $U(t) = E\left[\sum_{n=0}^\infty 1_{[0,t]}(S_n)\right] = \sum_{n=0}^\infty F^{n*}(t)$.

**Example 4.2.** Suppose that $X$ is an exponential random variable with parameter $\alpha$. Then

$$F(dx) = \alpha e^{-\alpha x} 1_{[0,\infty)}(x)$$

and hence

$$\hat{F}(\lambda) = \int_0^\infty e^{-\lambda x} \alpha e^{-\alpha x} dx = \frac{\alpha}{\alpha + \lambda}$$

**Example 4.3.** Suppose $Y$ has Gamma distribution with parameters $(n + 1)$ and $\alpha$, which we call an **Erlang distribution**. Suppose $Y$ has distribution $G$. Then,

$$G(dx) = \frac{\alpha(\alpha x)^n e^{-\alpha x}}{n!} 1_{[0,\infty)}(x)$$

and hence

$$\begin{aligned}\hat{G}(\lambda) &= \int_0^\infty e^{-\lambda x} \frac{\alpha(\alpha x)^n e^{-\alpha x}}{n!} dx \\ &= \alpha^{n+1} \int_0^\infty \frac{e^{-(\alpha+\lambda)x} x^n}{n!} \cdot \frac{(\alpha + \lambda)^{n+1}}{(\alpha + \lambda)^{n+1}} dx \\ &= \frac{\alpha^{n+1}}{(\alpha + \lambda)^{n+1}} \underbrace{\int_0^\infty \frac{e^{-(\alpha+\lambda)x} x^n (\alpha + \lambda)^{n+1}}{n!} dx}_{=1} \\ &= \left(\frac{\alpha}{\alpha + \lambda}\right)^{n+1}\end{aligned}$$

So the sum of $n + 1$ i.i.d. exponential r.vs with parameter $\alpha$ is Erlang with parameters $(n + 1)$ and $\alpha$.

**Definition 4.3.** Suppose that $S_0 = Y_0$ has distribution $G$ and $\{Y_n : n \geq 1\}$ has distribution $F$. Define

$$V(t) = \sum_{n=0}^{\infty} P(S_n \leq t) = \sum_{n=0}^{\infty} G * F^{(n-1)*}(t), F^{0*}(t) = 1_{[0,\infty)}(t)$$

*Remark* 4.4. Note that $\{N(t) \leq n\} = \{S_n > t\}$ from the monotonicity of $S_n$ and in general $S_{N(t)-1} \leq t < S_{N(t)}$. This will give us $\{N(t) = n\} = \{S_{n-1} \leq t < S_n\}$ and $\{N(t) = n\}$ only depends on $S_0, S_1, ..., S_n$.

**Theorem 4.1.** *For any $t \geq 0$,*

*1) $\sum_{n=0}^{\infty} \gamma^n F^{n*}(t) < \infty$ for $\gamma < 1/F(0)$.*

*2) The moment generating function of $N(t)$ exists $\implies$ all moments are finite and in particular $U(t)$.*

*Proof.* (See Resnik et al.) □

**Example 4.4.** Suppose $F$ is exponential where $F(dx) = \alpha e^{-\alpha x}\, dx$ for $x \geq 0$ and directly, we can compute

$$\begin{aligned} U(t) = \sum_{n=0}^{\infty} F^{n*}(t) &= F^{0*}(t) + \sum_{n=1}^{\infty} \int_0^t \frac{\alpha(\alpha u)^{n-1} e^{-\alpha u}}{(n-1)!} du \\ &= 1 + \int_0^t \alpha \underbrace{\sum_{n=1}^{\infty} \frac{(\alpha u)^{n-1} e^{-\alpha u}}{(n-1)!}}_{=1} du \\ &= 1 + \int_0^t \alpha du \\ &= 1 + \alpha t \end{aligned}$$

**Example 4.5.** Consider $F(dx) = xe^{-x}\, dx$ for $x \geq 0$. Consider

$$\begin{aligned} \left(\widehat{\sum_{n=1}^{\infty} F^{n*}}\right)(\lambda) &= \int_0^{\infty} e^{-\lambda x} \sum_{n=1}^{\infty} F^{n*}(dx) \\ &= \sum_{n=1}^{\infty} \int_0^{\infty} e^{-\lambda x} F^{n*}(dx) \\ &= \sum_{n=1}^{\infty} \hat{F}^{n*}(\lambda) \\ &= \sum_{n=1}^{\infty} \left[\hat{F}(\lambda)\right]^n \\ &= \sum_{n=1}^{\infty} \left(\frac{1}{1+\lambda}\right)^{2n} \\ &= \frac{1}{(1+\lambda)^2} \sum_{n=1}^{\infty} \left(\frac{1}{(1+\lambda)^2}\right)^{n-1} \\ &= \frac{1}{\lambda(\lambda+2)} \\ &= \frac{1}{2\lambda} - \frac{1}{2(\lambda+2)} \end{aligned}$$

Re-writing back to integrals, we get

$$\left(\widehat{\sum_{n=1}^{\infty} F^{n*}}\right)(\lambda) = \int_0^\infty \frac{1}{2}e^{-\lambda x}dx - \int_0^\infty \frac{1}{2}e^{-(\lambda+2)x}dx$$
$$= \int_0^\infty e^{-\lambda x}\left(\frac{1}{2} - \frac{1}{2}e^{-2x}\right)dx$$

and since this is equal to $\int_0^\infty e^{-\lambda x}\sum_{n=1}^\infty F^{n*}(dx)$, we have $\sum_{n=1}^\infty F^{n*}(dx) = \left(\frac{1}{2} - \frac{1}{2}e^{-2x}\right)dx$ and

$$U(t) = 1 + \int_0^t \left(\frac{1}{2} - \frac{1}{2}e^{-2x}\right)dx = \frac{3}{4} + \frac{1}{2}t + \frac{1}{4}e^{-2t}$$

**Theorem 4.2.** *Suppose that $\mu = E[Y_1] = \int_0^\infty xF(dx) < \infty$.*

*1) If $P(Y_0 < \infty) = 1$ then as $t \to \infty$ we have $N(t)/t \to 1/\mu$ almost surely.*

*2) Suppose that $\sigma^2 = Var(Y_1) < \infty$. Then as $t \to \infty$, $N(t)$ has a normal distribution with mean $t/\mu$ and variance $t\sigma^2/\mu^3$ and*

$$P\left(\frac{N(t) - t/\mu}{\sqrt{t\sigma^2/\mu^3}} < x\right) = N(0, 1, x)$$

*Proof.* 1) We can directly compute

$$\lim_{n\to\infty} \frac{S_n}{n} = \lim_{n\to\infty} \frac{Y_0 + Y_1 + ... + Y_n}{n}$$
$$= \lim_{n\to\infty}\left(\frac{Y_0}{n} + \frac{Y_1 + ... + Y_n}{n}\right)$$
$$= \mu \text{ a.s.}$$

from the CLT. Now $N(t)$ is non-decreasing in $t$. We need $N(t) \to \infty$ as $t \to \infty$ with probability 1. Since

$$\{N(t) > n\} = \{S_n \le t\}$$

then

$$P(N(t) > n) = G * F^{(n-1)*}(t) \to 1$$

Hence, we may use the fact that

$$S_{N(t)-1} \le t < S_{N(t)} \implies \frac{S_{N(t)-1}}{N(t)} \le \frac{t}{N(t)} < \frac{S_{N(t)}}{N(t)}$$
$$\implies \frac{S_{N(t)-1}}{N(t)} \cdot \frac{N(t)-1}{N(t)-1} \le \frac{t}{N(t)} < \frac{S_{N(t)}}{N(t)}$$
$$\implies \mu \le \lim_{t\to\infty} \frac{t}{N(t)} < \mu$$

and so $N(t)/t \to 1/\mu$.

2) We know that

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = N(0, 1, x)$$

from the CLT. Now,

$$P\left(\frac{N(t)-t/\mu}{\sqrt{\sigma^2 t/\mu^3}} \le x\right) = P\left(N(t) \le x\left(\sigma^2 t/\mu^3\right)^{1/2} + t/\mu\right)$$
$$= P\left(N(t) \le \underbrace{\left\lfloor x\left(\sigma^2 t/\mu^3\right)^{1/2} + t/\mu\right\rfloor}_{h(t)}\right)$$

so since $\{N(t) \le n\} = \{S_n > t\}$ then

$$P\left(\frac{N(t)-t/\mu}{\sqrt{\sigma^2 t/\mu^3}} \le x\right) = P(S_{h(t)} > t) = P\left(\frac{t-h(t)\mu}{\sigma\sqrt{h(t)}} > \frac{t-h(t)\mu}{\sigma\sqrt{h(t)}}\right)$$

We need $h(t) \to \infty$ and $[t-h(t)\mu]/[\sigma\sqrt{h(t)}] \to -x$. To get this, remark that

$$\lim_{t\to\infty} \frac{h(t)}{t/\mu} = 1 \implies h(t) \to \infty$$

and since

$$h(t) = x\left(\sigma t/\mu^3\right)^{1/2} + t/\mu + \varepsilon(t), |\varepsilon(t)| < 1$$

then

$$\frac{t-h(t)\mu}{\sigma\sqrt{h(t)}} = \frac{t-\mu\left(\sigma^2 t/\mu^3\right)^{1/2} x - t - \mu\varepsilon(t)}{\sigma\sqrt{h(t)}}$$
$$\to \frac{-\mu t^{1/2} x\sigma/\mu^{3/2}}{t^{1/2}/\mu^{3/2}}$$
$$\to -x$$

This gives us

$$P\left(\frac{N(t)-t/\mu}{\sqrt{\sigma^2 t/\mu^3}} \le x\right) = P(S_{h(t)} > t) \to N(0,1,x)$$

□

**Theorem 4.3.** *(Elementary Renewal Theorem) Let $\mu = E[Y_1] < \infty$ and $P(Y_0 < \infty) = 1$. Then,*

$$\lim_{t\to\infty} \frac{V(t)}{t} = \lim_{t\to} \frac{U(t)}{t} = \frac{1}{\mu}$$

*Proof.* We have

$$\frac{1}{\mu} = E\left[\lim_{t\to\infty} \frac{N(t)}{t}\right] \le \liminf_{t\to\infty} E\left[\frac{N(t)}{t}\right] = \liminf_{t\to\infty} \frac{U(t)}{t} = \liminf_{t\to\infty} \frac{V(t)}{t}$$

So define

$$Y_0^* = 0, Y_i^* = \min(Y_i, b), b > 0$$
$$S_0^* = 0, S_n^* = Y_0^* + ... + Y_n^*$$
$$N^*(t) = \sum_{n=0}^{\infty} 1_{[0,t]}(S_n^*)$$

where we have $S_n \ge S_n^*$ and $N^*(t) \ge N(t)$. Using **Wald's Lemma** which states that $E[S_{N(t)}] = E[N(t)]E[Y_1]$, we have

$$E[S^*_{N^*(t)}] = E[N^*(t)]E[Y_1^*]$$

and hence

$$\limsup_{t\to\infty}\frac{V(t)}{t}\le\limsup_{t\to\infty}\frac{V^*(t)}{t}=\limsup_{t\to\infty}\frac{E[S^*_{N^*(t)}]}{E[Y_1^*]}\cdot\frac{1}{t}=\limsup_{t\to\infty}\frac{E[S^*_{N^*(t)-1}+Y^*_{N^*(t)}]}{E[Y_1^*]}\cdot\frac{1}{t}$$

and from the bounds of $Y_i^*$ we have

$$\limsup_{t\to\infty}\frac{V(t)}{t}\le\limsup_{t\to\infty}\frac{t+b}{E[Y_1^*]t}=\frac{1}{E[Y_1^*]}=\frac{1}{E[\min(Y_1,b)]}$$

Since $\lim_{b\to\infty}E[\min(Y_1,b)]=E[Y_1]$ then

$$\frac{1}{\mu}\le\liminf_{t\to\infty}\frac{V(t)}{t}\le\limsup\frac{V(t)}{t}=\frac{1}{\mu}$$

□

**Renewal Reward Process**

Suppose we have a renewal sequence $\{S_n\}$ and suppose that at each epoch $S_n$we receive a random reward $R_n$. Suppose that $\{R_n : n\ge 1\}$ is a sequence of iid r.vs and define

$$R(t)=\sum_{i=0}^{\infty}R_i1(S_i\le t)=\sum_{i=1}^{N(t)-1}R_i$$

**Proposition 4.1.** *If* $E[|R_j|]<\infty$ *for all* $j=0,1,...$ *and* $E[Y_1]<\infty$ *with* $P(Y_0<\infty)=1$ *then*

$$\lim_{t\to\infty}\frac{R(t)}{t}=\frac{E[R_1]}{\mu}$$

*Proof.* We have

$$\lim_{t\to\infty}\frac{R(t)}{t}=\lim_{t\to\infty}\frac{\sum_{i=0}^{N(t)-1}R_i}{t}=\lim_{t\to\infty}\frac{\sum_{i=0}^{N(t)-1}R_i}{t}\cdot\frac{N(t)-1}{N(t)-1}=\frac{E[R_1]}{\mu}\text{ a.s.}$$

□

*Remark* 4.5. Suppose that $\{N(t):t\ge 0\}$ is independent of $\{R_n\}$. Then

$$\lim_{t\to\infty}\frac{E[R(t)]}{t}=\frac{E[R_1]}{\mu}$$

## 4.4 Renewal Equation

Consider the **renewal equation**

$$Z=z+F*Z\implies Z(t)=z(t)+\int_0^t Z(t-s)F(ds)$$

This is the case for $U(t)$ as follows:

$$\begin{aligned}U(t)&=\sum_{n=0}^{\infty}F^{n*}(t)\\&=F^{0*}(t)+\sum_{n=1}^{\infty}F^{n*}(t)\\&=F^{0*}(t)+F*\sum_{n=1}^{\infty}F^{(n-1)*}(t)\\U(t)&=F^{0*}(t)+F*U(t)\end{aligned}$$

**Example 4.6.** (Forward and Backward Recurrence Times) Define the **backward recurrence time** (age) $A(t)$ and **forward recurrence time** (excess life, residual life, etc.) $B(t)$ as

$$\begin{aligned} A(t) &= t - S_{N(t)-1} \\ B(t) &= S_{N(t)} - t \end{aligned}$$

**[Backward]** We have

$$P(A(t) \le x) = P(A(t) \le x, Y_1 > t) + P(A(t) \le x, Y_1 \le t)$$

The first term is

$$\begin{aligned} P(A(t) \le x, Y_1 > t) &= P(A(t) \le x | Y_1 < t) P(Y_1 > t) \\ &= 1_{[0,x]}(t)[1 - F(t)] \end{aligned}$$

and the second term is

$$\begin{aligned} P(A(t) \le x, Y_1 \le t) &= P(A(t) \le x, S_1 \le t) \\ &= P(A(t) \le x, N(t) \ge 2) \\ &= P(t - S_{N(t)-1} \le x, N(t) \ge 2) \\ &= \sum_{n=2}^{\infty} P(t - S_{n-1} \le x, N(t) = n) \\ &= \sum_{n=2}^{\infty} P(t - S_{n-1} \le x, S_{n-1} \le t < S_n) \\ &= \sum_{n=2}^{\infty} \int_0^t P(t - S_{n-1} \le x, S_{n-1} \le t < S_n | Y_1 = y) F(dy) \\ &= \sum_{n=2}^{\infty} \int_0^t P\left(t - \left(y + \sum_{i=2}^{n-1} Y_i\right) \le x, y + \sum_{i=2}^{n-1} Y_i \le t < y + \sum_{i=2}^{n} Y_i\right) F(dy) \\ &= \sum_{n=2}^{\infty} \int_0^t P\left(t - y - S_{n-2} \le x, y + S_{n-2} \le t < y + S_{n-1}\right) F(dy) \\ &= \sum_{n=2}^{\infty} \int_0^t P\left(t - y - S_{n-2} \le x, S_{n-2} \le t - y < S_{n-1}\right) F(dy) \\ &= \sum_{n=2}^{\infty} \int_0^t P\left(t - y - S_{N(t-y)-1} \le x, N(t-y) = n-1\right) F(dy) \\ &= \sum_{n=1}^{\infty} \int_0^t P\left(A(t-y) \le x, N(t-y) = n\right) F(dy) \\ &= \int_0^t P(A(t-y) \le x) F(dy) \end{aligned}$$

**[Forward]** We have

$$P(B(t) > x) = P(B(t) > x, S_1 > t) + P(B(t) > x, S_1 \le t)$$

The first part is

$$P(B(t) > x, S_1 > t) = P(S_1 > t + x) = 1 - F(t + x)$$

and the second part, using similar derivations for the the forward recurrence, is

$$P(B(t) > x, S_1 \le t) = \int_0^t P(B(t-y) > y) F(dy)$$

and hence

$$P(B(t) > x) = 1 - F(t + x) + \int_0^t P(B(t-y) > y) F(dy)$$

**Theorem 4.4.** *Suppose $Z(t) = 0$ for $t < 0$ and $z$ is locally bounded. Furthermore, assume that $F(0) < 1$. Then,*

*(i) A locally bounded solution of the renewal equation is*

$$U * z(t) = \int_0^t z(t-s)U(ds)$$

*(ii) There is no other locally bounded solution vanishing on $(-\infty, 0)$.*

*Proof.* (1) We will first show that $U * z$ is a locally bounded for $T > 0$. We have

$$\sup_{0 \le t \le T} U * z(t) = \sup_{0 \le t \le T} \int_0^t z(t-y)U(dy) \le \left( \sup_{0 \le s \le T} z(s) \right) \int_0^t U(dy) \le \left( \sup_{0 \le s \le T} z(s) \right) [U(t)]$$

Now

$$Z = z + F * Z \implies F * Z = Z - z$$

and hence

$$F * (U * z) = (F * U) * z = \left( F * \sum_{n=0}^{\infty} F^{n*} \right) * z = (U - F^{0*}) * z = U * z - z = Z - z$$

(2) Let $Z_1$ and $Z_2$ be two solutions that are locally bounded and vanishing on $(-\infty, 0)$. Define $H = Z_1 - Z_2$ and note that $H$ is also locally bounded. We then have

$$H = Z_1 - Z_2 = F * Z_1 - F * Z_2 = F * (Z_1 - Z_2) = F^{2*} * (Z_1 - Z_2) = ... = F^{n*} * (Z_1 - Z_2) = ...$$

and so

$$\begin{aligned} \sup_{0 \le t \le T} |H(t)| &= \sup_{0 \le t \le T} \left| \int_0^t \left( Z_1(t-y) - Z_2(t-y) \right) F^{n*}(dy) \right| \\ &\le \left| \sup_{0 \le s \le T} H(s) \right| F^{n*}(T) \end{aligned}$$

As $n \to \infty$ we have $\left| \sup_{0 \le s \le T} H(s) \right| F^{n*}(T) \to 0$. □

**Example 4.7.** Coming back to our forward and backward recurrence equations, recall that

$$P(A(t) \le x) = 1_{[0,x]}(t)[1 - F(t)] + \int_0^t P(A(t-y) \le x)F(dy)$$

$$P(B(t) > x) = [1 - F(t+x)] + \int_0^t P(B(t-y) > y)F(dy)$$

A locally bounded solution for the forward recurrence equation, using our previous theorem, is

$$P(A(t) \le x) = \int_0^t \left(1 - F(t-y)\right) 1_{[0,x]}(t-y)\, U(dy)$$

In the particular case of $F(dx) = \alpha e^{-\alpha x}\, dx, U(t) = 1 + \alpha t$ with $x \ge 0$, we have for the forward recurrence equation:

$$\begin{aligned} P(A(t) \le x) &= \int_0^t \left(1 - F(t-y)\right) 1_{[0,x]}(t-y)\, U(dy) \\ &= (1 - F(t)) + \int_0^t e^{-\alpha(t-y)} 1_{[0,x]}(t-y) \alpha dy \end{aligned}$$

If $t \leq x$ then

$$\begin{aligned} P(A(t) \leq x) &= (1 - F(t)) + \int_0^t \alpha e^{-\alpha(t-y)} dy \\ &= e^{-\alpha t} + e^{-\alpha t} e^{\alpha y} \Big|_{y=0}^{y=t} \\ &= 1 \end{aligned}$$

If $t > x$ then

$$\begin{aligned} P(A(t) \leq x) &= (1 - F(t)) + \int_{t-x}^t \alpha e^{-\alpha(t-y)} dy \\ &= e^{-\alpha t} + e^{-\alpha t} e^{\alpha y} \Big|_{y=t-x}^{y=t} \\ &= 1 - e^{-\alpha x} \end{aligned}$$

In summary,

$$P(A(t) \leq x) = \begin{cases} 1 & t \leq x \\ 1 - e^{-\alpha x} & t > x \end{cases}$$

In the case of the backward recurrence equation:

$$\begin{aligned} P(B(t) > x) &= \int_0^t (1 - F(t + x - y)) U(dy) \\ &= (1 - F(t + x)) + \int_0^t \alpha e^{-\alpha(t+x-y)} dy \\ &= e^{-\alpha(t+x)} + e^{-\alpha(t+x)} \int_0^t \alpha e^{\alpha y} dy \\ &= e^{-\alpha(t+x)} + e^{-\alpha(t+x)} e^{\alpha y} \Big|_0^t \\ &= e^{-\alpha(t+x)} + e^{-\alpha x} - e^{-\alpha(t+x)} \\ &= e^{-\alpha x} \end{aligned}$$

*Remark* 4.6. Observe that

$$\begin{aligned} F(dx) = \alpha e^{-\alpha x}, x \geq 0 \implies F^{n*}(dx) &= \frac{\alpha(\alpha x)^{n-1} e^{-\alpha x}}{(n-1)!} dx, x \geq 0 \\ \implies F^{n*}(x) &= \int_0^x \frac{\alpha(\alpha u)^{n-1} e^{-\alpha u}}{(n-1)!} du = 1 - \sum_{k=0}^{n-1} \frac{e^{-\alpha x}(\alpha x)^k}{k!} \end{aligned}$$

and since $\{N(t) = n + 1\}$ if and only if $\{S_n \leq t \leq S_{n+1}\}$ then,

$$\begin{aligned} P(N(t) = n + 1) &= P(S_n \leq t \leq S_{n+1}) \\ &= F^{n*}(t) - F^{(n+1)*}(t) \\ &= \sum_{k=0}^{n} \frac{e^{-\alpha t}(\alpha t)^k}{k!} - \sum_{k=0}^{n-1} \frac{e^{-\alpha t}(\alpha t)^k}{k!} \\ &= \frac{e^{-\alpha t}(\alpha t)^n}{n!} \end{aligned}$$

and so

$$\begin{aligned} U(t) &= \sum_{n=0}^{\infty}(1+n)\frac{e^{-\alpha t}(\alpha t)^n}{n!} \\ &= \sum_{n=0}^{\infty}\frac{e^{-\alpha t}(\alpha t)^n}{n!} + \sum_{n=0}^{\infty}\frac{ne^{-\alpha t}(\alpha t)^n}{n!} \\ &= 1 + \alpha t\sum_{n=1}^{\infty}\frac{e^{-\alpha t}(\alpha t)^{n-1}}{(n-1)!} \\ &= 1 + \alpha t \end{aligned}$$

**Theorem 4.5.** *(Blackwell's Theorem) If $V(t, t+a] = E[N(t+a)] - E[N(t)]$ then*

$$\frac{V(t, t+a]}{t} \to \frac{a}{\mu}$$

**Theorem 4.6.** *(Key Renewal Theorem) We have*

$$\lim_{t\to\infty} Z(t) = \lim_{t\to\infty} z * U(t) = \frac{1}{\mu}\int_0^{\infty} z(s)\, ds$$

**Example 4.8.** In our backward recurrence equation, we have

$$\begin{aligned} \lim_{t\to\infty} P(B(t) > x) &= \frac{1}{\mu}\int_0^{\infty}(1 - F(x+s))\, ds \\ &= \frac{1}{\mu}\int_x^{\infty}(1 - F(s))\, ds \\ &= 1 - F_0(x) \end{aligned}$$

and for our forward recurrence equation, we have

$$\begin{aligned} \lim_{t\to\infty} P(A(t) \le x) &= \frac{1}{\mu}\int_0^{\infty}(1 - F(s))1_{[0,x]}(s)\, ds \\ &= \frac{1}{\mu}\int_0^{x}(1 - F(s))\, ds \\ &= F_0(x) \end{aligned}$$

where $F_0$ is called the **equilibrium distribution**.

**Example 4.9.** If $F(dx) = \alpha e^{-\alpha x}\, dx$ for $x \ge 0$ then

$$1 - F_0(x) = \alpha\int_x^{\infty} e^{-\alpha x} dx = 1 - e^{\alpha x} = 1 - F(x)$$

The Laplace transform of $F_0$ is

$$\hat{F}_0(\lambda) = \int_0^{\infty} e^{-\lambda x}\frac{1}{\mu}(1 - F(x))\, dx = \frac{1}{\mu}\int_0^{\infty} e^{-\lambda x}(1 - F(x))\, dx$$

Since $\int_0^{\infty} e^{-\lambda x}F(x)\, dx = \hat{F}(\lambda)/\lambda$, then

$$\hat{F}_0(\lambda) = \frac{1}{\lambda\mu}(1 - \hat{F}(\lambda))$$

**Example 4.10.** Consider a delayed renewal process with $G = F_0$. We know that $V(t) = G * U(t) = G * \sum_{n=0}^{\infty} F^{n*}(t)$ and $\hat{V}(\lambda) = \hat{G}(\lambda)\hat{U}(\lambda)$. If $F(dx) = \alpha e^{-\alpha x}\, dx$ again, then

$$\hat{V}(\lambda) = \frac{(1 - \hat{F}(\lambda))}{\lambda\mu} \cdot \frac{1}{(1 - \hat{F}(\lambda))} = \frac{1}{\lambda\mu} \implies V(t) = \frac{t}{\mu}$$

Conversely, if $V(t) = t/\mu$ then

$$\hat{V}(\lambda) = \frac{1}{\lambda\mu} = \hat{G}(\lambda)\hat{U}(\lambda) = \frac{\hat{G}(\lambda)}{1-\hat{F}(\lambda)} \implies \hat{G}(\lambda) = \frac{1-\hat{F}(\lambda)}{\lambda\mu} \implies G = F_0$$

## 4.5 Direct Riemann Integrability

**Definition 4.4.** Suppose $z(t) = 0$ for $t < 0$ and $z(t) \geq 0$ for $t \geq 0$. Consider an interval $[0, a]$ and define for $k \geq 1$,

$$\begin{aligned}
\underline{m}_k(h) &= \inf_{(k-1)h \leq t < kh} z(t) \\
\underline{\sigma}(h) &= \sum_{k: kh \leq a} h\underline{m}_k(h) \\
\overline{m}_k(h) &= \sup_{(k-1)h \leq t < kh} z(t) \\
\overline{\sigma}(h) &= \sum_{k: kh \leq a} h\overline{m}_k(h)
\end{aligned}$$

Recall that a function $z$ is **Riemann integrable** if

$$\lim_{h\to\infty} \overline{\sigma}(h) = \underline{\sigma}(h) = 0$$

**Definition 4.5.** On the other hand $z$ is Riemann integrable on $[0, \infty)$ if $\lim_{a\to\infty} \int_0^a z(s)\, ds$ exists. Then,

$$\int_0^\infty z(s)\, ds = \lim_{a\to\infty} \int_0^a z(s)\, ds$$

For **direct Riemann integrability** define $\underline{m}_k(h)$ and $\overline{m}_k(h)$ as above and define

$$\begin{aligned}
\underline{\sigma}(h) &= \sum_{k=1}^{\infty} h\underline{m}_k(h) \\
\overline{\sigma}(h) &= \sum_{k=1}^{\infty} h\overline{m}_k(h)
\end{aligned}$$

A function $z$ is directly Riemann integrable if $\overline{\sigma}(h) < \infty$ for all $h$ and

$$\lim_{h\to\infty} \overline{\sigma}(h) - \underline{\sigma}(h) = 0$$

**Example 4.11.** See Resnik p. 232 for an example of a (triangle) function which is Riemann integrable but not direct Riemann integrable.

*Remark* 4.7. Here are some facts from Resnik:

(1) If $z$ has a compact support then Riemann integrability is the same as direct Riemann integrability.

(2) If $z$ is directly Riemann integrable then it is Riemann integrable.

(3) If $z \geq 0$ and $z$ is non-increasing then $z$ is directly Riemann integrable if and only if it is Riemann integrable.

(4) If $z$ is Riemann integrable on $[0, a]$ for all $a > 0$ and $\sigma(1) < \infty$ then $z$ is directly Riemann integrable.

(5) If $z$ is Riemann integrable on $[0, \infty)$ and $z \leq g$ where $g$ is directly Riemann integrable then $z$ is directly Riemann integrable.

**Theorem 4.7.** *Suppose that $F(\infty) = 1$ and $F(0) < 1$. Define*

$$\mu = \int_0^\infty xF(dx), F_0(x) = \frac{1}{\mu}\int_0^x (1 - F(y))\, dy$$

*The following are equivalent:*

*(i) The Blackwell Theorem: If* $G(\infty) = 1$ *then*

$$\lim_{t\to\infty} V(t, t+b] = \frac{b}{\mu} \text{ for } b > 0$$

*(ii) The Key Renewal Theorem: Suppose* $z(t)$ *is directly Riemann integrable. Then*

$$\lim_{t\to\infty} U * z(t) = \frac{1}{\mu}\int_0^\infty z(s)\, ds$$

*(iii) Suppose that* $G(\infty) = 1$. *Then*

$$\lim_{t\to\infty} P(B(t) \le x) = F_0(x)$$

*(iv) Suppose that* $G(\infty) = 1$. *Then*

$$\lim_{t\to\infty} P(A(t) \le x) = F_0(x)$$

*Proof.* We will start with the equivalence of **(iii) and (iv)**. Note that

$$P(B(t) \le x) = P(N(t, t+x] \ge 1) = P(A(t+x) \le x)$$

and as $t \to \infty$ the probabilities are equal. For **(ii)** $\Longrightarrow$ **(iv)** we have

$$\begin{aligned} P(A(t) \le x) &= 1_{[0,x]}(t)[1 - F(t)] + \int_0^t P(A(t-y) \le x) F(dy) \\ &= \int_0^t (1 - F(t-y))\, 1_{[0,x]}(t-y)\, U(dy) \end{aligned}$$

and from (ii) we have

$$\lim_{t\to\infty} P(A(t) \le x) = \frac{1}{\mu}\int_0^\infty (1 - F(s)) 1_{[0,x]}(s)\, ds = \frac{1}{\mu}\int_0^x (1 - F(s))\, ds = F_0(x)$$

If we have a delayed renewal process, then

$$P(A(t) \le x) = P(A(t) \le x, S_0 > t) + \int_0^t P(A(t-y) \le x) G(dy)$$

and since

$$P(A(t) \le x, S_0 > t) \le P(S_0 > t) \overset{t\to\infty}{\Longrightarrow} \lim_{t\to\infty} P(A(t) \le x, S_0 > t) \le 0$$

Define $f_t(y) = P(A(t-y) \le x) 1_{[0,t]}(y)$. If $t, y > 0$ and $f_t(y) \le 1$ then

$$\lim_{t\to\infty} P(A(t-y) \le x) 1_{[0,t]}(y) = F_0(x)$$

Hence,

$$\begin{aligned} \lim_{t\to\infty} P(A(t) \le x) &= \lim_{t\to\infty} \int_0^t P(A(t-y) \le x) G(dy) \\ &= \lim_{t\to\infty} \int_0^\infty P(A(t-y) \le x) 1_{[0,t]}(y) G(dy) \\ &= \int_0^\infty \lim_{t\to\infty} P(A(t-y) \le x) 1_{[0,t]}(y) G(dy) \\ &= \int_0^\infty F_0(x) G(dy) \\ &= F_0(x) \end{aligned}$$

**(iii) and (iv)** have an equivalent formulation and clearly **(ii)** $\Longrightarrow$ **(iv), (ii)** $\Longrightarrow$ **(iii)**. We will next show that **(iii)** $\Longrightarrow$ **(i)**. We first have

$$\begin{aligned}
V(t,t+b] &= \int_t^{t+b} E[N(t+b-S_{N(t)})|S_{N(t)}=x]G_t(d[x-t]), P(B(t)\le x)=G_t(x) \\
&= \int_t^{t+b} E[N(t+b-x)]G_t(d[x-t]) \\
&= \int_t^{t+b} U(t+b-x)G_t(d[x-t]) \\
&= \int_0^b U(b-x)G_t(dx) \\
V(t,t+b] &= \int_0^b G_t(b-x)U(dx)
\end{aligned}$$

*The reasoning is that we are counting from the first renewal after time $t$ which will randomly depend on $S_{N(t)}$. However, $S_{N(t)}$ can be derived from $B(t)$ if given $t$ and so if we count from that first renewal, from the regenerative property of renewal processes this is the same as counting from a pure renewal process betweeen $t\in[0,t+b-S_{N(t)}]$.* Hence

$$\begin{aligned}
\lim_{t\to\infty} V(t,t+b] &= \lim_{t\to\infty}\int_0^b G_t(b-x)U(dx) \\
&= \int_0^b \lim_{t\to\infty} G_t(b-x)U(dx) \\
&= \int_0^b F_0(b-x)U(dx) \\
&= F_0 * U(b) \\
&= \frac{b}{\mu}
\end{aligned}$$

□

**Lemma 4.1.** *If $F(b)<1$ then $U(t-b,t]\le 1/(1-F(b))$ for all $t\ge b$. Thus,*

$$\sup_{t\ge 0} U(t,t+b] \le \frac{1}{1-F(b)} = c(b) < \infty$$

*Proof.* Since

$$U = F^{0*} + F*U \implies U(t) - F*U(t) = F^{0*}(t)$$

then

$$\begin{aligned}
1 = \int_0^t (1-F(t-s))U(ds) &\ge \int_{t-b}^t (1-F(t-s))U(ds) \\
&\ge \int_{t-b}^t (1-F(b))U(ds) \\
&= (1-F(b))\int_{t-b}^t U(ds) \\
&= (1-F(b))U(t-b,t]
\end{aligned}$$

□

**Theorem 4.8.** *Blackwell's Theorem implies the Key Renewal Theorem.*

*Proof.* Assume first that $z(t) = 1_{[(n-1)h,nh]}(t)$ and note that

$$z(t-s)=1 \iff (n-1)h \le t-s \le nh \iff t-nh \le s \le t-(n-1)h$$

and hence from Blackwell's Theorem,

$$\begin{aligned}\lim_{t\to\infty} U * z(t) &= \lim_{t\to\infty}\int_0^t z(t-s)U(ds)\\ &= \lim_{t\to\infty}\int_{t-nh}^{t-(n-1)h} U(ds)\\ &= \lim_{t\to\infty} U(t-nh, t-(n-1)h]\\ &= \frac{h}{\mu}\end{aligned}$$

Now since $\int_0^\infty z(t)\ dt = \int_{(n-1)h}^{nh} dt = h$, we have

$$\lim_{t\to\infty} U * z(t) = \frac{1}{\mu}\int_0^\infty z(t)\ dt = \frac{h}{\mu}$$

Now suppose that $z(t) = \sum_{n=1}^\infty c_n 1_{[(n-1)h,nh]}(t)$ where $c_n > 0$ and $\sum_{n=1}^\infty c_n < \infty$. From Blackwell's Theorem, for each $n$

$$U(t-nh, t-(n-1)h] \to \frac{h}{\mu} \text{ as } t\to\infty$$

Furthermore, from the previous lemma,

$$\sup_{t,n} U(t-nh, t-(n-1)h] \le c(h) < \infty$$

and hence

$$\begin{aligned}U * z(t) &= \int_0^t z(t-s)U(ds)\\ &= \int_0^t \sum_{n=1}^\infty c_n 1_{[(n-1)h,nh]}(t-s)U(ds)\\ &= \sum_{n=1}^\infty \int_{t-nh}^{t-(n-1)h} c_n U(ds)\\ &= \sum_{n=1}^\infty c_n U(t-nh, t-(n-1)h]\end{aligned}$$

Taking limits,

$$\begin{aligned}\lim_{t\to\infty} U * z(t) &= \lim_{t\to\infty}\sum_{n=1}^\infty c_n U(t-nh, t-(n-1)h]\\ &= \sum_{n=1}^\infty \lim_{t\to\infty} c_n U(t-nh, t-(n-1)h]\\ &= \sum_{n=1}^\infty \frac{c_n h}{\mu}\\ &= \frac{1}{\mu}\sum_{n=1}^\infty c_n h\\ &= \frac{1}{\mu}\int_0^\infty z(t)\ dt\end{aligned}$$

using the same reasoning as the simple $z(t)$ case. That is, $\int_0^\infty z(t)\ dt = \sum_{n=1}^\infty c_n \int_{(n-1)h}^{nh} = \sum_{n=1}^\infty c_n h$.

Next, assume that $z$ is a directly Riemann integrable function with

$$\overline{z}(t) = \sum_{n=1}^{\infty} \overline{m}_n(h) 1_{[(n-1)h, nh]}$$
$$\underline{z}(t) = \sum_{n=1}^{\infty} \underline{m}_n(h) 1_{[(n-1)h, nh]}$$

where

$$\underline{m}_n(h) = \inf_{(n-1)h \leq t < nh} z(t)$$
$$\overline{m}_n(h) = \sup_{(n-1)h \leq t < nh} z(t)$$

From direct Riemann integrability,

$$\sum_{n=1}^{\infty} \underline{m}_n(h) \leq \sum_{n=1}^{\infty} \overline{m}_n(h) < \infty$$

and from the previous step,

$$\lim_{t\to\infty} U * \overline{z}(t) = \frac{1}{\mu} \sum_{n=1}^{\infty} \overline{m}_n(h) h = \frac{\overline{\sigma}(h)}{\mu}$$
$$\lim_{t\to\infty} U * \underline{z}(t) = \frac{1}{\mu} \sum_{n=1}^{\infty} \underline{m}_n(h) h = \frac{\underline{\sigma}(h)}{\mu}$$

Since for any $h$, we have

$$\frac{\underline{\sigma}(h)}{\mu} = \liminf_{t\to\infty} U * \underline{z}(t) \leq \liminf_{t\to\infty} U * z(t) \leq \limsup_{t\to\infty} U * z(t) \leq \limsup_{t\to\infty} U * \overline{z}(t) = \frac{\overline{\sigma}(h)}{\mu}$$

then taking $h \to 0$ we have

$$\lim_{h\to 0} [\overline{\sigma}(h) - \underline{\sigma}(h)] = 0$$

and we are done. □

**Example 4.12.** ($\lim_{t\to\infty}[U(t) - t/\mu]$) Recall that $t/\mu = F_0 * U(t)$ and so

$$Z(t) = U(t) - \frac{t}{\mu} = U(t) - F_0 * U(t)$$
$$= (1 - F_0) * U(t)$$

From the key renewal theorem,

$$\lim_{t\to\infty} Z(t) = \frac{1}{\mu} \int_0^{\infty} (1 - F_0(t))\, dt$$

if $F_0$ is directly Riemann integrable. This is the case if and only if $\int_0^\infty \frac{u^2}{2} F(du) < \infty$.

*Proof.* (Blackwell's Theorem) We want to prove

$$V(t, t+a] \to \frac{a}{\mu} \text{ as } t \to \infty$$

Let us define $g(a) = \lim_{t\to\infty} V(t, t+a] = \lim_{t\to\infty}(V(t+a) - V(t))$ and note that

$$V(t+a+b) - V(t) = V(t+a+b) - V(t+a) + V(t+a) - V(t)$$
$$\implies g(a+b) = \lim_{t\to\infty} [V(t+a+b) - V(t+a)] + \lim_{t\to\infty} [V(t+a) - V(t)]$$
$$\implies g(a+b) = g(a) + g(b)$$

Suppose that $g(a) = ca, c > 0 \implies \lim_{n\to\infty} X_n = g(1) = c$. Define $\{X_n : n \geq 1\}$ such that $X_n = V(n) - V(n-1)$ for all $n \geq 1$ and remark that $\sum_{j=1}^n X_j = V(n) - V(0)$ from telescoping. Now,

$$c = \lim_{n\to\infty} \frac{\sum_{j=1}^n X_j}{n} = \lim_{n\to\infty} \frac{V(n) - V(0)}{n} = \frac{1}{\mu}$$

and $g(a) = a/\mu$. □

## 4.6 Regenerative Processes

**Definition 4.6.** Consider a stochastic process $\{X(t) : t \geq 0\}$ and let $\{S_n : n \geq 0\}$ be a renewal process. Then the process $\{X(t) : t \geq 0\}$ is called a **regenerative** process with regeneration points $\{S_n\}$ if

$$(X(S_n + t_i), i = 1, 2, ..., k) \stackrel{d}{=} (X(t_i), i = 1, 2, ..., k)$$

*Remark* 4.8. Suppose that $S_0 = 0$ and let $Z(t) = P(X(t) \in A)$. Then,

$$\begin{aligned} Z(t) &= P(X(t) \in A, S_1 > t) + P(X(t) \in A, S_1 \leq t) \\ &= K(t, A) + \int_0^t P(X(t) \in A | S_1 = s) F(dx) \\ &= K(t, A) + \int_0^t Z(t - s) F(ds) \end{aligned}$$

From the renewal equation, we get that $Z(t) = K(\cdot, A) * U(t)$.

**Theorem 4.9.** *(Smith's Theorem) Suppose $\{X(t)\}$ is a regenerative process with state space $E$. For fixed $A$, assume that $K(t, A)$ is Riemann integrable. Set $\mu \in E[S_1]$ and $S_0 = 0$.*

*a) If $\mu < \infty$, then*

$$\begin{aligned} \lim_{t\to\infty} P(X(t) \in A) &= \frac{1}{\mu} \int_0^\infty K(s, A)\, ds \\ &= \frac{1}{\mu} E\left[\int_0^{S_1} 1[X(s) \in A]\, ds\right] \\ &= \frac{E\left[\text{time spent in } A \text{ in a cycle}\right]}{E\left[\text{cycle length}\right]} \end{aligned}$$

*b) If $\mu = \infty$, then $\lim_{t\to\infty} P(X(t) \in A) = 0$.*

*Note that $K(t, A) \leq P(S_1 > t) = 1 - F(t)$.*

**Example 4.13.** Consider an M/G/1 queue. That is, the arrival process is Poisson and there is a single server whose service time has a general distribution. Assume that the arrival rate is $\alpha$. Let $X(t)$ be the number of customers in the system at time $t$. Suppose we would like to compute $\lim_{t\to\infty} P(X(t) = 0)$.

To do this, suppose that between the epochs $S_n$ and $S_{n+1}$ we have a busy period where at least one customer arrives. If $E[BP]$ = expected length of the busy period, then

$$\lim_{t\to\infty} P(X(t) = 0) = \frac{\frac{1}{\alpha}}{\frac{1}{\alpha} + E[BP]}$$

**Example 4.14.** (Alternating Renewal Processes) Consider a system that can be in one of two states: on or off. Initially it is on and it remains on for a time of $Z_1$ and then goes off and remains off for a period of $Y_1$. Then it remains on for an amount of time $Z_2$ and off for an amount of time $Y_2$, so on and so forth. Suppose that $\{(Z_n, Y_n) : n \geq 1\}$ is an i.i.d. sequence.

Define

$$X(t) = \begin{cases} 1 & \text{if the system is on at time } t \\ 0 & \text{otherwise} \end{cases}$$

Then,

$$\lim_{t\to\infty} P(X(t) = 1) = \frac{E[Z_1]}{E[Z_1] + E[Y_1]}$$

## 4.7 Poisson Random Variable

**Theorem 4.10.** *(Law of Small Numbers) If $n \to \infty$ and $p \to 0$ in such a way that $np \to \alpha$, then the binomial distribution with parameters $(n, p)$ converges to the Poisson distribution. That is for each $k = 0, 1, ...$ we have*

$$\binom{n}{k} p^k (1-p)^{n-k} \to \frac{\alpha^k e^{-\alpha}}{k!}$$

**Proposition 4.2.** *Let $T_n$ be a sequence of geometric random variables with parameters $p_n$ where $P(T_n > k) = (1-p_n)^k$ for $k = 0, 1, ....$ If $np_n \to \alpha$ as $n \to \infty$ then $T_n/n$ converges in distribution to the exponential distribution with parameter $\alpha$.*

*Proof.* Set $\alpha_n \to np_n$. Then, $\alpha_n \to \alpha$ as $n \to \infty$ and $p_n = \alpha_n/n$. So $P(T_n > k) = \left(1 - \frac{\alpha_n}{n}\right)^k$ and

$$\lim_{n\to\infty} P\left(\frac{T_n}{n} > t\right) = \lim_{n\to\infty} P(T_n > nt) = \lim_{n\to\infty} \left(1 - \frac{\alpha_n}{n}\right)^{\lceil nt \rceil} = e^{-\alpha t}$$

□

**Proposition 4.3.** *If $X_1, X_2, ..., X_n$ are independent Poisson random variables with $E[X_i] = \alpha_i$ then $\sum_{i=1}^n X_i$ is a Poisson random variable with mean $\alpha_1 + \alpha_2 + ... + \alpha_n$.*

**Fact 4.1.** *For a Poisson random variable,*

$$P(X = k) = \frac{e^{-\alpha}\alpha^k}{k!}, E[X] = \alpha, Var(X) = \alpha$$

**Theorem 4.11.** *Suppose that $N$ is a Poisson random variable with parameter $\alpha$ and $X_1, X_2, ...$ are i.i.d. Bernoulli random variables with parameter $p$ independent of $N$. Let $S_n = \sum_{i=1}^n X_i$. Then, $S_N$ is a Poisson random variable with mean $\alpha p$.*

*Proof.* We have

$$\begin{aligned}
P(S_N = k) &= \sum_{n=k}^{\infty} P(S_N = k | N = n) P(N = n) \\
&= \sum_{n=k}^{\infty} P(X_1 + X_2 + ... + X_n = k) P(N = n) \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{e^{-\alpha}\alpha^n}{n!} \\
&= \frac{p^k}{k!} e^{-\alpha} \alpha^k \sum_{n=k}^{\infty} \frac{((1-p)\alpha)^{n-k}}{(n-k)!} \\
&= \frac{p^k}{k!} e^{-\alpha} \alpha^k e^{(1-p)\alpha} \\
&= \frac{(\alpha p)^k e^{-\alpha p}}{k!}
\end{aligned}$$

□

**Theorem 4.12.** *(Generalized Thinning Theorem) Suppose $N$ is a Poisson random variable with parameter $\alpha$ and the $X_1, X_2, ...$ are i.i.d. multinomial random variables with parameters $(p_1, p_2, ..., p_m)$. That is,*

$$P(X_i = k) = p_k \text{ for each } k = 1, 2, ..., m$$

*Then the random variables $N_1, N_2, ..., N_m$ defined as*

$$N_k = \sum_{i=1}^{N} 1\{X_i = k\}$$

*are i.i.d. Poisson random variables with $E[N_k] = \alpha p_k$.*

**Definition 4.7.** A **point process** on the timeline $[0, \infty)$ is a mapping $J \mapsto N_j = N(j)$ that assigns to each subset $J \subset [0, \infty)$ a non-negative integer value random variable $N_j$ in such a way that if $J_1, J_2, ...$ are pairwise disjoint then

$$N(\cup_i J_i) = \sum_i N(J_i)$$

We will interchangeably use $N(t) = N([0, t])$.

**Definition 4.8.** (Poisson process) A **Poisson point process** of intensity $\alpha > 0$ is a point process $N(J)$ with the following properties:

a) If $J_1, J_2, ...$ are non-overlapping intervals of $[0, \infty)$ then the random variables $N(J_1), N(J_2), ...$ are mutually independent. (*Independent Increments*)

b) For every interval $J$,

$$P(N(J) = k) = \frac{e^{-\alpha|J|}(\alpha|J|)^k}{k!},\ k = 0, 1, ...$$

where $|J|$ is the length of the interval $J$.

**Theorem 4.13.** *Define $0 = S_0 \leq S_1 \leq S_2 \leq ...$ as the successive times that the process $N(t)$ has jumps. Define the interarrival times as $Y_n = S_n - S_{n-1}$.*

*(a) The interarrival times $Y_1, Y_2, ...$ of a Poisson process with rate $\alpha$ are i.i.d. exponential random variables with mean $1/\alpha$.*

*(b) Conversely let $X_1, X_2, ...$be i.i.d. exponential random variables with mean $1/\alpha$ and define*

$$N(t) = \max\left\{n : \sum_{i=1}^{n} X_i \leq t\right\}$$

*Then $\{N(t) : t \geq 0\}$ is a Poisson process with rate $\alpha$.*

*Proof.* (a) We have

$$P(S_1 > t) = P(Y_1 > t) = P(N(t) = 0) = e^{-\alpha t}$$

and

$$\begin{aligned} P(Y_1 > t, Y_2 > s) &= P(S_1 > t, Y_2 > s) \\ &= \int_t^\infty P(S_1 > t, Y_2 > s | S_1 = u) F(du) \\ &= \int_t^\infty P(N(u, s+u] = 0) F(du) \\ &= \int_t^\infty e^{-\alpha s} F(du) \\ &= e^{-\alpha s} \int_t^\infty F(du) \\ &= e^{-\alpha(t+s)} \end{aligned}$$

□

**Theorem 4.14.** *For each $m \geq 1$, let $\{X_r^m : r \in N/m\}$ be a **Bernoulli process** indexed by the integer multiples of $1/m$ with probability of success $p_m$. Let $\{N^m(t)\}$ be the corresponding counting process that is*

$$N^m(t) = \sum_{r \leq t} X_r^m$$

*If* $\lim_{m\to\infty} mp_m = \alpha > 0$ *then for any finite set of points* $0 = t_0 < t_1 < ... < t_n$

$$(N^m(t_1), N^m(t_2), ..., N^m(t_n)) \xrightarrow{D} (N(t_1), N(t_2), ..., N(t_n))$$

*where* $\xrightarrow{D}$ *means convergence in distribution.*

*Proof.* Define

$$\Delta_k^m = (N^m(t_k) - N^m(t_{k-1})), \Delta_k = (N(t_k) - N(t_{k-1}))$$

From the Law of small numbers,

$$(\Delta_1^m, \Delta_2^m, ..., \Delta_n^m) \xrightarrow{D} (\Delta_1, \Delta_2, ..., \Delta_n)$$

□

*Proof.* [cont. from the previous Theorem] (a) The interarrival (interoccurence) times of a Bernoulli process is geometric, but in this case the interarrival times are scaled by $1/m$. Thus, from the previous part of the proof, the interarrival times of the limit process are exponential. [**This uses the implicit relationship between the occurrence times and the interoccurrence times**]

[cont. from the previous Theorem] (b) Recall that $\{X_i\}$ is a sequence of independent exponentially distributed random variables with parameter $\alpha$. We have $S_0 = 0, S_n = \sum_{i=1}^n Y_i$. Set $T_n = \sum_{i=1}^n X_i$. Now,

$$N^{(Y)}(t) \sim (T_1, T_2, ..., T_n) \stackrel{D}{=} (S_1, S_2, ..., S_n) \sim N(t)$$

The result then holds for the corresponding counting process. □

**Definition 4.9.** The (stationary) counting process $\{N(t) : t \geq 0\}$ is said to be a Poisson process with intensity $\alpha > 0$ if:

(i) the process has independent increments

(ii) $P(N(h) = 1) = \alpha h + o(h)$

(iii) $P(N(h) \geq 2) = o(h)$

Recall that a function $f$ is $o(h)$ if $\lim_{h\to\infty}(f(h)/h) = 0$.

*Remark* 4.9. To see the previous definition, let us first show that $P(N(t) = 0) = e^{-\alpha t}$. We have

$$\begin{aligned} P(N(t+h) = 0) &= P(N(t) = 0, N(t+h) - N(t) = 0) \\ &= P(N(t) = 0)P(N(t+h) - N(t) = 0) \\ &= P(N(t) = 0)\,(1 - P(N(t+h) - N(t) = 1) - P(N(t+h) - N(t) \geq 2)) \\ P(N(t+h) = 0) &= P(N(t) = 0)(1 - \alpha h + o(h)) \end{aligned}$$

and so

$$\begin{aligned} P'(N(t) = 0) = \lim_{h\to 0} \frac{P(N(t+h) = 0) - P(N(t) = 0)}{h} &= \lim_{h\to 0} \frac{\alpha h P(N(t) = 0)}{h} + \lim_{h\to 0} \frac{o(h)P(N(t) = 0)}{h} \\ &= -\alpha P(N(t) = 0) \end{aligned}$$

Hence, $P(N(t) = 0) = Ce^{-\alpha t}$. At $t = 0, C = 1$ and so $P(N(t) = 0) = e^{-\alpha t}$. Next, for $n \geq 1$,

$$\begin{aligned} P(N(t+h) = n) =& P(N(t) = n, N(t+h) - N(t) = 0) + \\ & P(N(t) = n-1, N(t+h) - N(t) = 1) + \\ & \sum_{k\geq 2}^{\infty} P(N(t) = n-k, N(t+h) - N(t) = k) \end{aligned}$$

and note that

$$\sum_{k\geq 2}^{\infty} P(N(t) = n-k, N(t+h) - N(t) = k) \leq \sum_{k\geq 2}^{\infty} P(N(t+h) - N(t) = k) = P(N(t+h) - N(t) \geq 2) = o(h)$$

Hence,

$$\begin{aligned} P(N(t+h)=n) &= P(N(t)=n)(1-\alpha h+o(h)) + P(N(t)=n-1)(\alpha h+o(h)) + o(h) \\ &= P(N(t)=n)(1-\alpha h) + P(N(t)=n-1)(\alpha h) + o(h) \end{aligned}$$

and thus

$$\begin{aligned} P'(N(t)=n) = \lim_{h\to 0}\frac{P(N(t+h)=0)-P(N(t)=0)}{h} &= \lim_{h\to 0}\frac{\alpha h P(N(t)=n)}{h} + \lim_{h\to 0}\frac{\alpha h P(N(t)=n-1)}{h} \\ &= -\alpha P(N(t)=n) + \alpha P(N(t)=n-1) \end{aligned}$$

This gives us the equation

$$e^{\alpha t}\left[P'(N(t)=n)+\alpha P(N(t)=n)\right] = \frac{d}{dt}\left(e^{\alpha t}P(N(t)=n)\right) = \alpha e^{-\alpha t}P(N(t)=n-1)$$

For $n=1$,

$$\frac{d}{dt}(e^{-\alpha t}P(N(t)=1)) = \alpha \implies e^{-\alpha t}P(N(t)=1) = \alpha t + C \implies P(N(t)=1) = \alpha t e^{\alpha t} + Ce^{\alpha t}$$

At $t=0, C=0$ and $P(N(t)=1)=e^{-\alpha t}(\alpha t)$. Now assume that $P(N(t)=n-1)=(e^{-\alpha t}(\alpha t)^{n-1})/(n-1)!$. We have

$$\frac{d}{dt}(e^{-\alpha t}P(N(t)=1)) = \frac{\alpha(\alpha t)^{-n-1}}{(n-1)!} \implies P(N(t)=n) = \frac{\alpha^n t^n e^{-\alpha t}}{n!} + Ce^{-\alpha t}$$

and at $t=0, C=0$ to get

$$P(N(t)=n) = \frac{e^{-\alpha t}(\alpha t)^n}{n!}$$

**Proposition 4.4.** *Given that $N[0,1]=k$, the $k$ points are uniformly distributed on the unit interval $[0,1]$, that is for any partition $J_1, J_2, ..., J_m$ of $[0,1]$ into non-overlapping intervals*

$$P(N(J_i)=k_i, i=1,2,...,m|N[0,1]=k) = \frac{k!}{k_1!k_2!...k_m!}\prod_{i=1}^m |J_i|^{k_i}$$

*for all non-negative integers $k_1,...,k_m$ with $\sum_{i=1}^m k_i = k$.*

*Proof.* Picky $\sum_{i=1}^m k_i = k$ and directly evaluate:

$$\begin{aligned} &P(N(J_i)=k_i, i=1,2,...,m|N[0,1]=k) \\ =&\frac{P(N(J_i)=k_i, i=1,2,...,m, N[0,1]=k)}{P(N[0,1]=k)} \\ =&\frac{P(N(J_i)=k_i, i=1,2,...,m)}{P(N[0,1]=k)} \\ =&\frac{\prod_{i=1}^m \frac{(\alpha|J_i|)e^{-\alpha|J_i|}}{k_i!}}{\frac{e^{-\alpha}\alpha^k}{k!}} \\ =&\frac{\prod_{i=1}^m \frac{|J_i|^{k_i}}{k_i!}}{\frac{1}{k!}} \\ =&\frac{k!}{k_1!k_2!...k_m!}\prod_{i=1}^m |J_i|^{k_i} \end{aligned}$$

□

**Proposition 4.5.** *Let $S_1, S_2, ...$ be the arrival times of a Poisson process $\{N(t) : t \geq 0\}$ with rate $\alpha$. Then conditional on the event that $N[0,t]=k$, the variables $S_1, S_2, ..., S_k$ are distributed in the same manner as the order statistics of i.i.d. uniform $[0,t]$ random variables.*

**Proposition 4.6.** *Suppose that each event of a Poisson process is classified as a type I process with probability $p(s)$ when the event happens at time $s$ and type II with probability $1-p(s)$. Suppose $\{N(t) : t \geq 0\}$ is a Poisson process with rate $\alpha$. If $N_1(t)$ and $N_2(t)$ represent the type I and type II events, respectively by time $t$, then $N_1(t)$ and $N_2(t)$ are independent Poisson random variables with means $\lambda_1 = \alpha \int_0^t p(s)\, ds$ and $\lambda_2 = \alpha \int_0^t (1-p(s))\, ds$.*

*Proof.* We need to show

$$P(N_1(t) = n, N_2(t) = m) = \frac{e^{-\lambda_1}(\lambda_1)^n}{n!} \cdot \frac{e^{-\lambda_2}(\lambda_2)^m}{m!}$$

Directly we have

$$\begin{aligned} &P(N_1(t) = n, N_2(t) = m) \\ =&\sum_{k=0}^{\infty} P(N_1(t) = n, N_2(t) = m | N(t) = k) P(N(t) = k) \\ =&P(N_1(t) = n, N_2(t) = m | N(t) = n+m) P(N(t) = n+m) \end{aligned}$$

Since

$$\begin{aligned} &P(\text{an arrival of type I in } [0,t] \mid \text{an arrival in } [0,t]) \\ =&\int_0^t \underbrace{P(\text{a type I event} \mid \text{an event at time } s)}_{p(s)} \underbrace{P(\text{an event time } s \mid \text{an event in } [0,t])}_{1/t}\, ds \\ =&\frac{1}{t}\int_0^t p(s)\, ds \end{aligned}$$

and similarly

$$P(\text{an arrival of type I in } [0,t] \mid \text{an arrival in } [0,t]) = \frac{1}{t}\int_0^t (1-p(s))\, ds$$

then we have

$$\begin{aligned} &P(N_1(t) = n, N_2(t) = m | N(t) = n+m) \\ =&\binom{n+m}{n} \left(\frac{1}{t}\int_0^t p(s)\, ds\right)^n \left(\frac{1}{t}\int_0^t (1-p(s))\, ds\right)^m \end{aligned}$$

and

$$\begin{aligned} &P(N_1(t) = n, N_2(t) = m) \\ =&P(N_1(t) = n, N_2(t) = m | N(t) = n+m) P(N(t) = n+m) \\ =&\frac{(n+m)!}{n!m!} \left(\frac{1}{t}\int_0^t p(s)\, ds\right)^n \left(\frac{1}{t}\int_0^t (1-p(s))\, ds\right)^m \frac{e^{-\alpha t}(\alpha t)^{n+m}}{(n+m)!} \\ =&\frac{\left(\alpha \int_0^t p(s)\, ds\right)^n e^{-\alpha t\left(\frac{1}{t}\int_0^t p(s)\, ds\right)}}{n!} \cdot \frac{\left(\alpha \int_0^t (1-p(s))\, ds\right)^m e^{-\alpha t\left(\frac{1}{t}\int_0^t (1-p(s))\, ds\right)}}{m!} \\ =&\frac{e^{-\lambda_1}(\lambda_1)^n}{n!} \cdot \frac{e^{-\lambda_2}(\lambda_2)^m}{m!} \end{aligned}$$

with the fact that

$$\frac{1}{t}\int_0^t p(s)\, ds + \frac{1}{t}\int_0^t (1-p(s))\, ds = \frac{1}{t}\int_0^t ds = \frac{t}{t} = 1$$

□

**Definition 4.10.** Let $m(t) = \int_0^t \alpha(s)\, ds$. The counting process $\{N(t) : t \geq 0\}$ is said to be a non-stationary (non-homogeneous) Poisson process with intensity function $\alpha(t), t \geq 0$ if

(i) $P(N(0) = 0) = 1$.

(ii) $\{N(t) : t \geq 0\}$ has independent increments.

(iii) We have

$$P(N(t+s) - N(t) = n) = \frac{e^{-(m(t+s)-m(t))}(m(t+s) - m(t))^n}{n!}, n \geq 0$$

**Definition 4.11.** The counting process $\{N(t) : t \geq 0\}$ is said to be a non-stationary (non-homogeneous) Poisson process with intensity function $\alpha(t), t \geq 0$ if

(i) $P(N(0) = 0) = 1$.

(ii) $\{N(t) : t \geq 0\}$ has independent increments.

(iii) $P(N(t+h) - N(t) = 1) = \alpha(t)h + o(h)$

(iv) $P(N(t+h) - N(t) \geq 1) = o(h)$

**Example 4.15.** For a M/G/$\infty$ queue, we have $\alpha(t) = \alpha \int_0^t p(s)\ ds$ and mean number of active services at time $t$ equal to $\alpha \int_0^t \int_{t-s}^{\infty} G(dy)\ ds$ where $p(s) = \int_{t-s}^{\infty} G(dy)\ ds$.