

Introduction

Schedule:

- (I) Basic properties of continuous-time Markov Chains
- (II) Qualitative properties of continuous time Markov Chains
- (III) Queueing theory
- (IV) Renewal theory
- (V) Spatial Poisson processes

1 Some basic aspects of continuous-time Markov Chains

Definition. A sequence of random variables is called a *stochastic process* or *process*. The process $X = (X_n)_{n \geq 1}$ is called a discrete-time Markov Chain with state space I if for all $x_0, x_1, \dots, x_n \in I$

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}).$$

If $\mathbb{P}(X_{n+1} = y | X_n = x)$ is independent of n , the chain is called *time-homogeneous*. We then write $P = (P_{x,y})_{x,y \in I}$ for the *transition matrix* where $P_{x,y} = \mathbb{P}(X_1 = y | X_0 = x)$. The data associated to every time-homogeneous Markov Chain is the transition matrix P and the initial distribution μ , i.e $\mathbb{P}(X_0 = x_0) = \mu(x_0)$.

From now on:

- I denotes a countable (or finite) state space.
- $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which all the relevant random variables are defined.

Definition. $X = (X(t) : t \geq 0)$ is a (right-continuous) continuous-time random process with values in I if

- (a) for all $t \geq 0$, $X(t) = X_t$ is a random variable such that $X(t) : \Omega \rightarrow I$;
- (b) for all $\omega \in \Omega$, $t \mapsto X_t(\omega)$ is right-continuous (right-continuous sample path).
In our case this means for all $\omega \in \Omega$, for all $t \geq 0$, there exists $\varepsilon > 0$ (depending on ω, t) such that

$$X_t(\omega) = X_s(\omega) \quad \forall s \in [t, t + \varepsilon].$$

Fact. A right-continuous random process is defined by its finite-dimensional distributions

$$\mathbb{P}(X_{t_0}=i, X_{t_1}=i_1, \dots, X_{t_n}=i_{t_n}), \quad n \geq 0, \quad t_k \geq 0, \quad i_k \in I.$$

For every $\omega \in \Omega$, the path $t \mapsto X_t(\omega)$ of a right-continuous process stays constant for a while. So there are 3 possibilities:

- (i) The path makes infinitely many jumps overall but only finitely many in a given interval $[0, t]$.
- (ii) The path makes finitely many jumps & then gets absorbed in some state.
- (iii) The path makes infinitely jumps in a finite time interval. After the ‘explosion time’ ζ , the process starts up again.

Write $J_0 = 0, J_1, J_2, \dots$ for the jump times and S_1, S_2, \dots for the holding times, defined by

$$J_0 = 0, \quad J_{n+1} = \inf\{t \geq J_n : X_t \neq X_{J_n}\},$$

$$S_n = \begin{cases} J_n - J_{n-1} & J_{n-1} < \infty \\ \infty & \text{otherwise} \end{cases}.$$

By right-continuity, $S_n > 0$ for all n . If $J_{n+1} = \infty$ for some n , we define $X_\infty = X_{J_n}$ as the final value, otherwise X_∞ is not defined. The explosion time ζ is defined by

$$\zeta = \sup(J_n) = \sum_{n=1}^{\infty} S_n.$$

We are not going to consider what happens to a chain after explosion. We thus set $X_t = \infty$ for all $t \geq \zeta$ (adjoining a new state ‘ ∞ ’). We call such a chain *minimal*.

Definition. We define the *jump chain* Y_n of $(X_t)_{t \geq 0}$ by setting $Y_n = X_{J_n}$ for all n .

Definition. A right-continuous random process $X = (X_t)_{t \geq 0}$ has the Markov property (and is called a continuous-time markov chain) if for all $i_1, i_2, \dots, i_n \in I$ and $0 \leq t_1 < t_2 < \dots < t_n$,

$$\mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \dots, X_{t_0} = i_0) = \mathbb{P}(X_{t_n} = i_n | X_{t_0} = i_0).$$

Remark. For all $h > 0$, $Y_n = X(hn)$ defines a discrete-time Markov Chain.

Definition. The transition probabilities are $P_{ij}(s, t) = \mathbb{P}(X_t = j | X_s = i)$, $s \leq t$, $i, j \in I$. It is called *time-homogeneous* if it depends on $t - s$ only, i.e

$$P_{ij}(s, t) = P_{i,j}(0, t - s).$$

In this case we just write $P_{ij}(t - s)$. As in the case of discrete time, a (time-homogeneous) Markov process is characterised by

- 1. Its initial distribution $\lambda_i = \mathbb{P}(X_0 = i)$, $i \in I$;
- 2. Its *family of transition matrices* $(P(t))_{t \geq 0} = (P_{ij}(t))_{t \geq 0}$.

The family $(P(t))_{t \geq 0}$ is called the *transition subgroup* of the MC.

A (time-homogeneous) Markov process is characterised by

- its initial distribution;
- its transition subgroup $(P(t))_{t \geq 0}$

$$(P(t))_{t \geq 0} = (P(t))_{\substack{i, j \in I \\ t \geq 0}} = (\mathbb{P}(X_t = j | X_0 = i))_{\substack{i, j \in I \\ t \geq 0}}$$

It is easy to see that

- $P(0)$ is the identity
- $P(t)$ is a stochastic matrix for all t (i.e rows sum to 1)
- $P(t+s) = P(t)P(s) \forall s, t$ (Chapman-Kolmogorov equation)

$$\begin{aligned} P_{xz}(t+s) &= \mathbb{P}(X_{t+s} = z | X_0 = x) \\ &= \sum_{y \in I} \mathbb{P}(X_{t+s} = z | X_0 = x, X_t = y) \mathbb{P}(X_t = y | X_0 = x) \\ &= \sum_{y \in I} \mathbb{P}(X_s = z | X_0 = y) \mathbb{P}(X_t = y | X_0 = x) \\ &= \sum_{y \in I} P_{yz}(s) P_{xy}(t) = P_{x \cdot}(t) P_{\cdot z}(s) \end{aligned}$$

Holding times

Let X be a (right-continuous continuous-time time-homogeneous) Markov Chain on a countable state-space I .

Suppose X starts from $x \in I$. Question: how long does X stay in the state x ?

Definition. We call S_x the *holding time* at state x ($S_x > 0$ by right-continuity).

Let $s, t \geq 0$. Then

$$\begin{aligned} \mathbb{P}(S_x > t+s | S_x > s) &= \mathbb{P}(X_u = x \forall u \in [0, t+s] | X_u = x \forall u \in [0, s]) \\ &= \mathbb{P}(X_u = x \forall u \in [s, t+s] | X_u = x \forall u \in [0, s]) \\ &= \mathbb{P}(X_u = x \forall u \in [s, t+s] | X_s = x) \\ &= \mathbb{P}(X_u = x \forall u \in [0, t] | X_0 = x) \\ &= \mathbb{P}(S_x > t). \end{aligned}$$

Thus S_x has the memoryless property.

By the next theorem, we will get that S_x has the exponential distribution, say with parameter q_x .

Theorem 1.1 (Memoryless property). *Let S be a positive random variable. Then S has the memoryless property, i.e. $\mathbb{P}(S > t + s | S > s) = \mathbb{P}(S > t)$ for all $s, t \geq 0$ if and only if S has the exponential distribution.*

Proof. It is easy to see the exponential distribution is memoryless. So we prove the other direction. Set $F(t) = \mathbb{P}(S > t)$. Then $F(s + t) = F(s)F(t)$ for all $s, t \geq 0$.

Since S is a positive random variable, there exists $n \in \mathbb{N}$ large such that $F(1/n) = \mathbb{P}(S > 1/n) > 0$. Then $F(1) = F(1/n)^n > 0$. So we can set $F(1) = e^{-\lambda}$ for some $\lambda \geq 0$.

For $k \in \mathbb{N}$, $F(k) = F(1)^k = e^{-\lambda k}$. For p/q rational, $F(p/q) = F(1/q)^p = (F(1/q)^q)^{p/q} = F(1)^{p/q} = e^{-\lambda \frac{p}{q}}$.

For any $t \geq 0$, for any $r, s \in \mathbb{Q}$ such that $r \leq t \leq s$, since F is decreasing

$$e^{-\lambda s} = F(s) \leq F(t) \leq F(r) = e^{-\lambda r}.$$

So taking sequences of rationals approaching t , we have $F(t) = e^{-\lambda t}$. \square

Poisson Process'

We are now going to look at the simplest (and most important) example of continuous time Markov Chains - the Poisson process.

Definition. Suppose S_1, S_2, \dots are iid random variables with $S_1 \sim \text{Exp}(\lambda)$. Define the *jump times* $J_0 = 0, J_1 = S_1, J_n = S_1 + \dots + S_n$ for all n , and set $X_t = i$ if $J_i \leq t < J_{i+1}$. Then $I = \{0, 1, 2, \dots\}$ and note that X is right-continuous and increasing. X is called a *Poisson process* of parameter/intensity λ . We sometimes refer to the jump times $(J_i)_{i \geq 1}$ as the *points* of the Poisson process, then X = number of points in $[0, t]$.

Theorem 1.2 (Markov property). *Let $(X_t)_{t \geq 0}$ be a Poisson process of intensity λ . Then for all $s \geq 0$, the process $(X_{s+t} - X_s)_{t \geq 0}$ is also a Poisson process of intensity λ , and is independent of $(X_t)_{0 \leq t \leq s}$.*

Proof. Set $Y_t = X_{t+s} - X_s$ for all $t \geq 0$. Let $i \in \{0, 1, 2, \dots\}$ and condition on $\{X_s = i\}$. Then the jump times for the process Y are $J_{n+1} - s, J_{n+2} - s, \dots$ and the holding times are

$$\begin{aligned} T_1 &= J_{n+1} - s = S_{i+1} - (s - J_i) \\ T_2 &= S_{i+2} \\ T_3 &= S_{i+3} \\ &\vdots \end{aligned}$$

Since $\{X_s = i\} = \{J_i \leq s\} \cap \{s < J_{i+1}\} = \{J_i \leq s\} \cap \{S_{i+1} > s - J_i\}$, conditional on $\{X_s = i\}$, by the memoryless property of the exponential distribution (and

independence of S_{i+1} and J_i) we see that $T_1 \sim \text{Exp}(\lambda)$. Moreover the times J_j , $j \geq 2$ are independent of S_k , $k \leq i$ and hence independent of $(X_r)_{r \leq s}$, and they have iid $\text{Exp}(\lambda)$ distribution. Thus $((X_{s+t} - X_s))_{t \geq 0}$ is a Poisson process of parameter λ and is independent of $(X_t)_{0 \leq t \leq s}$. \square

Similar to this, one can show the Strong Markov property for a Poisson process of parameter λ . Recall a random variable $T \in [0, \infty]$ is called a *stopping time* if for all t , the event $\{T \leq t\}$ depends only on $(X_s)_{s \leq t}$.

Theorem 1.3 (Strong Markov property). *Let $(X_t)_{t \geq 0}$ be a Poisson process of parameter λ and T a stopping time. Then conditional on $T < \infty$, the process $(X_{T+t} - X_T)_{t \geq 0}$ is a Poisson process of parameter λ and independent of $(X_s)_{s \leq T}$.*

Theorem 1.4. Let $(X_t)_{t \geq 0}$ be an increasing right-continuous process taking values in $\{0, 1, 2, \dots\}$ with $X_0 = 0$. Let $\lambda > 0$. Then the following are equivalent

- (a) The holding times S_1, S_2, \dots are iid $\text{Exp}(\lambda)$ and the jump chain is given by $Y_n = n$ (i.e X is a poisson process of intensity λ)
- (b) (Infinitesimal def) X has independent increments and as $h \downarrow 0$ uniformly in t we have

$$\mathbb{P}(X_{t+h} - X_t = 1) = \lambda h + o(h),$$

$$\mathbb{P}(X_{t+h} - X_t = 0) = 1 - \lambda h + o(h).$$

- (c) X has independent and stationary increments and for all $t \geq 0$, $X_t \sim \text{Poi}(\lambda t)$.

Proof. First we show (a) \Rightarrow (b). If (a) holds, then by the Markov property, the increments are independent and stationary $((X_{t+s} - X_s)_{t \geq 0} \stackrel{d}{=} (X_t - X_0)_{t \geq 0})$. Using stationarity we have (uniformly in t) as $h \rightarrow 0$,

$$\mathbb{P}(X_{t+h} - X_t = 0) = \mathbb{P}(X_h = 0) = \mathbb{P}(S_1 > h) = e^{-\lambda h} = 1 - \lambda h + o(h),$$

$$\mathbb{P}(X_{t+h} - X_t \geq 1) = \mathbb{P}(X_h \geq 1) = \mathbb{P}(S_1 \leq h) = 1 - e^{-\lambda h} = \lambda h + o(h),$$

$$\begin{aligned} \mathbb{P}(X_{t+h} - X_t \geq 2) &= \mathbb{P}(X_h \geq 2) = \mathbb{P}(S_1 + S_2 \leq h) \\ &\leq \mathbb{P}(S_1 \leq h, S_2 \leq h) \\ &= \mathbb{P}(S_1 \leq h)^2 \\ &= (\lambda h + o(h))^2 = o(h). \end{aligned}$$

Now we show (b) \Rightarrow (c). If X satisfies (b), then $(X_{t+s} - X_s)_{t \geq 0}$ also satisfies (b). So X has independent and stationary increments. Now set $p_j(t) = \mathbb{P}(X_t = j)$. Then since increments are independent and X is increasing,

$$\begin{aligned} p_j(t+h) &= \mathbb{P}(X_{t+h} = j) = \sum_{i=0}^j \mathbb{P}(X_t = j-i) \mathbb{P}(X_{t+h} - X_t = i) \\ &= p_j(t)(1 - \lambda h + o(h)) + p_{j-1}(t)(\lambda h + o(h)) + o(h). \end{aligned}$$

Thus, $\frac{p_j(t+h) - p_j(t)}{h} = -\lambda p_j(t) + \lambda p_{j-1}(t) + o(1)$. Setting $s = t + h$, we get

$$\frac{p_j(s) - p_j(s-h)}{h} = -\lambda p_j(s-h) + \lambda p_{j-1}(s-h) + o(1).$$

In particular, $p_j(t)$ is continuous and differentiable with

$$p'_j(t) = -\lambda p_j(t) + \lambda p_{j-1}(t).$$

Differentiating

$$(e^{\lambda t} p(t))' = \lambda e^{\lambda t} p_j(t) + e^{\lambda t} p'_j(t) = \lambda e^{\lambda t} p_{j-1}(t).$$

For $j = 0$ we have $p_0(t + h) = p_0(t)(1 - \lambda h + o(h))$, i.e. $p'_0(t) = -\lambda p_0(t)$ so $p_0(t) = e^{-\lambda t}$. Thus

$$p'_1(t) = -\lambda p_1(t) + \lambda e^{-\lambda t}, \text{ i.e. } p_1(t) = \lambda t e^{-\lambda t}.$$

And by induction

$$p_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!},$$

i.e. $X_t \sim \text{Poi}(\lambda t)$.

Finally we show (c) \Rightarrow (a). We know X has independent stationary increments, We have for $t_1 \leq \dots \leq t_k$, $n_1 \leq \dots \leq n_k$,

$$\begin{aligned} & \mathbb{P}(X_{t_1} = n_1, \dots, X_{t_k} = n_k) \\ &= \mathbb{P}(X_{t_1} = n_1) \underbrace{\mathbb{P}(X_{t_2} - X_{t_1} = n_2 - n_1)}_{\sim \text{Poi}(\lambda(t_2 - t_1))} \dots \underbrace{\mathbb{P}(X_{t_k} - X_{t_{k-1}} = n_k - n_{k-1})}_{\sim \text{Poi}(\lambda(t_k - t_{k-1}))}. \end{aligned}$$

So (c) determines the finite-dimensional distributions (f.d.d) of a right-continuous process X , hence (c) determines X . So (c) \Rightarrow (a).

Question: can we show (a) \Rightarrow (c) directly? Indeed note

$$\begin{aligned} \mathbb{P}(X_t = n) &= \mathbb{P}(S_1 + \dots + S_n \leq t < S_1 + \dots + S_{n+1}) \\ &= \mathbb{P}(S_1 + \dots + S_n \leq t) - \mathbb{P}(S_1 + \dots + S_{n+1} \leq t) \\ &= \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} dx - \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^n}{n!} dx \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!} \text{ (integration by parts).} \end{aligned}$$

□

Theorem 1.5 (Superposition). *Let X and Y be two independent Poisson processes with parameters λ and μ respectively. Then $(Z_t)_{t \geq 0} = (X_t + Y_t)_{t \geq 0}$ is a Poisson process with parameter $\lambda + \mu$.*

Proof. We use (c) from the previous theorem. So Z has stationary independent increments. Also $Z_t \sim \text{Poi}(\lambda t + \mu t)$. □

Theorem 1.6 (Thinning). *Let X be a Poisson process with parameter λ . Let $(Z_i)_{i \geq 1}$ be a sequence of iid Bernoulli(p) random variables. Let Y be a Poisson process with values in $\{0, \dots\}$ which jumps at time t if and only if X_t jumps at time t and $Z_{X_t} = 1$.*

In other words, we keep every point of X with probability p independently. Then Y is another Poisson process, with parameter λp and $X - Y$ is an independent Poisson process with parameter $\lambda(1 - p)$.

Proof. We shall use the infinitesimal definition. The independence of increments for Y is clear. Since $\mathbb{P}(X_{t+h} - X_t \geq 2) = o(h)$, we have

$$\mathbb{P}(Y_{t+h} - Y_t = 1) = p\mathbb{P}(X_{t+h} - X_t = 1) + o(h) = p\lambda h + o(h),$$

$$\begin{aligned}\mathbb{P}(Y_{t+h} - Y_t = 0) &= \mathbb{P}(X_{t+h} - X_t = 0) + (1-p)\mathbb{P}(X_{t+h} - X_t = 1) + o(h) \\ &= 1 - \lambda h + (1-p)(\lambda h + o(h)) + o(h) \\ &= 1 - \lambda p h + o(h).\end{aligned}$$

Hence Y is Poisson of parameter λp . Clearly $X - Y$ is a thinning of X with Bernoulli parameter $1 - p$, so $X - Y$ is Poisson of parameter $\lambda(1 - p)$.

Now we show Y and $X - Y$ are independent. It is enough to show that the f.d.d of Y and $X - Y$ are independent, i.e if $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$, $n_1 \leq \dots \leq n_k$ and $m_1 \leq \dots \leq m_k$, then we want to prove

$$\begin{aligned}\mathbb{P}(Y_{t_1} = n_1, \dots, Y_{t_k} = n_k, X_{t_1} - Y_{t_1} = m_1, \dots, X_{t_k} - Y_{t_k} = m_k) \\ = \mathbb{P}(X_{t_1} = n_1, \dots, Y_{t_k} = n_k) \mathbb{P}(X_{t_1} - Y_{t_1} = m_1, \dots, X_{t_k} - Y_{t_k} = m_k).\end{aligned}$$

We will only show this for fixed t ($k = 1$) the general case follows similarly using independence of increments. We have

$$\begin{aligned}\mathbb{P}(Y_t = n, X_t - Y_t = m) &= \mathbb{P}(X_t = m + n, Y_t = n) \\ &= \mathbb{P}(X_t = m + n) \mathbb{P}(Y_t = n | X_t = m + n) \\ &= e^{-\lambda t} \frac{(\lambda t)^{m+n}}{(m+n)!} \binom{m+n}{n} p^n (1-p)^m \\ &= e^{-\lambda t p} \frac{(\lambda t p)^n}{n!} e^{-\lambda t(1-p)} \frac{(\lambda t(1-p))^m}{m!} \\ &= \mathbb{P}(X_t = n) \mathbb{P}(X_t - Y_t = m),\end{aligned}$$

as required. □

Theorem 1.7. *Let X be a Poisson Process. Conditional on the event $(X_t = n)$, the jump times J_1, J_2, \dots, J_n are distributed as the order statistics of n iid $U[0, t]$ random variables. That is, they have joint density*

$$f(t_1, \dots, t_n) = \frac{n!}{t^n} \mathbb{1}(0 \leq t_1 \leq \dots \leq t_n \leq t).$$

Proof. Since S_1, S_2, \dots are iid $\text{Exp}(\lambda)$, the joint density of (S_1, \dots, S_{n+1}) is

$$\lambda^{n+1} e^{-\lambda(S_1 + \dots + S_{n+1})} \mathbb{1}(S_i \geq 0 \text{ for all } i).$$

Then the jump times $J_1 = S_1, J_2 = S_1 + S_2, \dots, J_{n+1} = S_1 + \dots + S_{n+1}$ have joint density

$$g(t_1, \dots, t_{n+1}) = \lambda^{n+1} e^{-\lambda t_{n+1}} \mathbb{1}(0 \leq t_1 \leq t_2 \leq \dots \leq t_{n+1}).$$

(Noting the Jacobian of the transformation is 1.) Now take $A \subseteq \mathbb{R}^n$ so

$$\mathbb{P}((J_1, \dots, J_n) \in A | X_t = n) = \frac{\mathbb{P}((J_1, \dots, J_n) \in A | X_t = n)}{\mathbb{P}(X_t = n)}.$$

Note

$$\begin{aligned} & \mathbb{P}((J_1, \dots, J_n) \in A, X_t = n) \\ &= \mathbb{P}((J_1, \dots, J_n) \in A, J_n \leq t < J_{n+1}) \\ &= \int_{(t_1, \dots, t_{n+1}) \in A \times \mathbb{R}} g(t_1, \dots, t_n) \mathbb{1}(t_{n+1} \geq t \geq t_n) dt_1 \dots dt_{n+1} \\ &= \int_A \int_t^\infty \lambda^{n+1} e^{-\lambda t_{n+1}} \mathbb{1}(0 \leq t_1 \leq \dots \leq t_n \leq t) dt_{n+1} dt_1 \dots dt_n \\ &= \int_A \lambda^n e^{-\lambda t} \mathbb{1}(0 \leq t_1 \leq \dots \leq t_n \leq t) dt_1 \dots dt_n. \end{aligned}$$

Then we get

$$\mathbb{P}((J_1, \dots, J_n) \in A | X_t = n) = \int_A \frac{n!}{t^n} \mathbb{1}(0 \leq t_1 \leq \dots \leq t_n \leq t) dt_1 \dots dt_n.$$

As required. \square

Now we look at a generalisation of a Poisson Process: called a Birth Process. For a Poisson Process, the rate of going from i to $i+1$ is λ . For a Birth Process, this is q_i (can depend on i). More precisely:

Definition (Birth Process). For each i , let $S_i = \text{Exp}(q_i)$ with S_1, S_2, \dots independent. Set $J_i = S_1 + \dots + S_i$ and $X_t = i$ if $J_i \leq t < J_{i+1}$. Then X is called a *Birth Process*.

We have some special cases:

1. Simple birth process: when $q_i = \lambda i$ for $i = 1, 2, \dots$;
2. Poisson Process $q_i = \lambda$ for all i .

Motivation for Simple Birth Process (SBP): at time 0 there is only one ‘individual’ i.e $X_0 = 1$. Each individual has an exponential clock of parameter λ independently. Then if there are i individuals, the first clock rings after $\text{Exp}(\lambda i)$ time, and we jump from i to $i + 1$ individuals. Indeed, by the memoryless property, the process begins afresh after each jump.

Proposition 1.8. *Let $(T_k)_{k \geq 1}$ be a sequence of independent random variables with $T_K \sim \text{Exp}(q_k)$ and $\sum_k q_k < \infty$. Let $T = \inf_k T_k$. Then*

- (a) $T \sim \text{Exp}(\sum_k q_k)$
- (b) *The infimum is attained at a point T_K almost surely, and*

$$\mathbb{P}(K = n) = \frac{q_n}{\sum_k q_k}.$$

- (c) T and K are independent.

Proof. See example sheet. □

The main difference between a Poisson Process and a Birth Process is that there is the possibility of explosion in the Birth Process. Recall explosion occurs when $\zeta := \sum_n S_n < \infty$.

Proposition 1.9. *Let X be a Birth Process with rates q_i and $X_0 = 1$. Then*

1. *If $\sum_{i=1}^{\infty} \frac{1}{q_i} < \infty$, then X is explosive, i.e $\mathbb{P}(\zeta < \infty) = 1$;*
2. *If $\sum_{i=1}^{\infty} \frac{1}{q_i} = \infty$, then X is non-explosive, i.e $\mathbb{P}(\zeta = \infty) = 1$.*

Remark. This shows the SBP (as well as the PP) is non-explosive.

Proof.

1. If $\sum_n \frac{1}{q_n} < \infty$, then

$$\mathbb{E}[\zeta] = \mathbb{E} \left[\sum_n S_n \right] = \sum_n \mathbb{E} S_n = \sum_n \frac{1}{q_n} < \infty.$$

Where we have swapped summation and expectation by the MCT (monotone convergence theorem). Thus $\zeta = \sum_n S_n < \infty$ almost surely.

2. If $\sum_n \frac{1}{q_n} = \infty$, then $\prod_n \left(1 + \frac{1}{q_n}\right) \geq 1 + \sum_n \frac{1}{q_n} = \infty$. Then

$$\begin{aligned}
 \mathbb{E}[e^{-\zeta}] &= \mathbb{E}\left[e^{-\sum_{n=1}^{\infty} S_n}\right] \\
 &= \lim_{n \rightarrow \infty} \mathbb{E}\left[e^{-\sum_{i=1}^n S_i}\right] && \text{(MCT)} \\
 &= \lim_{n \rightarrow \infty} \prod_{i=1}^n \mathbb{E}[e^{-S_i}] && \text{(independence)} \\
 &\leq \lim_{n \rightarrow \infty} \prod_{i=1}^n \frac{1}{1 + 1/q_i} = 0.
 \end{aligned}$$

Since $e^{-\zeta} \geq 0$, since $\mathbb{E}(e^{-\zeta}) = 0$ we have $e^{-\zeta} = 0$ almsot surely, i.e $\mathbb{P}(\zeta = \infty) = 1$.

□

Theorem 1.10 (Markov Property). *Let X be a BP with parameters (q_i) . Conditional on $X_s = i$, the process $(X_{s+t})_{t \geq 0}$ is a birth process with rates $(q_j)_{j \geq i}$ starting from i , and independent of $(X_r)_{r \leq s}$.*

Proof. As in the Poisson Process case. \square

Theorem 1.11. *Let X be an increasing right-continuous process with values in $\{1, 2, \dots\} \cup \{\infty\}$. Let $0 \leq q_j < \infty$ for all $j \geq 0$. Then the following are equivalent:*

1. (jump chain/holding time definition) conditional on $X_s = i$, the holding times S_1, S_2, \dots are independent exponentials with rates q_i, q_{i+1}, \dots respectively and the jump chain is given $Y_n = i + n$ for all n .
2. (infinitesimal definition) for all $t, h \geq 0$, conditional on $X_t = i$, the process $(X_{t+h})_{h \geq 0}$ is independent of $(X_s)_{s \leq t}$ and as $h \rightarrow 0$, uniformly in t we have

$$\mathbb{P}(X_{t+h} = i | X_t = i) = 1 - q_i h + o(h),$$

$$\mathbb{P}(X_{t+h} = i + 1 | X_t = i) = q_i h + o(h).$$

3. (transition probability definition) for all $n = 0, 1, 2, \dots$ and all times $0 \leq t_0 \leq t_1 \leq \dots \leq t_{n+1}$, and all states i_0, i_1, \dots, i_{n+1} ,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_0 = i_0, \dots, X_{t_n} = i_n) = p_{i_n, i_{n+1}}(t_{n+1} - t_n),$$

where $(p_{ij}(t) : i, j = 0, 1, 2, \dots)$ is the unique solution to the equation (called Kolmogorov's forward equation)

$$p'_{ij}(t) = q_{j-1} p_{i, j-1}(t) - q_j p_{i, j}(t). \quad (*)$$

(as in the Poisson Process, $p_{ij}(t+h) = p_{i, j-1}(t) q_j h + p_{i, j}(t) (1 - q_j h) + o(h)$.)

Existence and uniqueness of a solution in (3) follow since for $i = j$ $p'_{i, i}(t) = -q_i p_{i, i}(t)$ and $p_{i, i}(0) = 1$, so $p_{i, i}(t) = e^{-q_i t}$. Then by induction, if the unique solution for $p_{i, j}(t)$ exists, then plug into (*) to see there exists a unique solution for $p_{i, j+1}(t)$.

Also note that we can write the equation in matrix form:

$$P'(t) = P(t)Q, \text{ where } Q = \begin{pmatrix} -q_1 & q_1 & 0 & \dots \\ 0 & -q_2 & q_2 & \dots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Q-matrix and construction of Markov Processes

Definition. $Q = (q_{ij})_{i, j \in I}$ is called a Q-matrix if

- (a) $-\infty < q_{ii} \leq 0$ for all $i \in I$;

(b) $0 \leq q_{ij} < \infty$ for all $i, j \in I$ with $i \neq j$;

(c) $\sum_{j \in I} q_{ij} = 0$ for all $i \in I$.

Write $q_i = -q_{ii} = \sum_{j \notin I} q_{ij}$ for all $i \in I$.

Given a Q -matrix Q , we define a jump matrix P as follows. For $x \neq y$ with $q_x \neq 0$, set $p_{xy} = \frac{q_{xy}}{q_x}$ and $p_{xx} = 0$. If $q_x = 0$, set $p_{xy} = \mathbb{1}(x = y)$.

Example.

$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix} \implies P = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 2/3 & 1/3 & 0 \end{pmatrix}.$$

Definition. Let Q be a Q -matrix and λ a probability measure on the state space I . Then a (minimal) random process X is a *Markov process* with initial distribution λ and infinitesimal generator Q if

- (a) The jump chain $Y_n = X_{J_n}$ is a discrete time Markov chain starting from $Y_0 \sim \lambda$ with transition matrix P .
- (b) Conditional on Y_0, Y_1, \dots, Y_n , the holding times S_1, \dots, S_{n+1} are independent with $S_i \sim \text{Exp}(q_{Y_{i-1}})$ for $i = 1, \dots, n+1$.

We write $X \sim \text{Markov}(\lambda, Q)$.

Example. Birth-Processes are $\text{Markov}(\lambda, Q)$ with $I = \mathbb{N}$ and

$$Q = \begin{pmatrix} -q_1 & q_1 & 0 & \dots \\ 0 & -q_2 & q_2 & \dots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \text{ and } P = \begin{pmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

And jump chain $Y_n = Y_0 + n$.

We have multiple constructions of a Markov(λ, Q) process:

Construction 1:

- $(Y_n)_{n \geq 1}$ is a discrete-time Markov chain, $Y_0 \sim \lambda$ & transition matrix P .
- $(T_i)_{i \geq 1}$ iid Exp(1) random variables, independent of Y and set $S_n = \frac{T_n}{q_{Y_{n-1}}}$ and $J_n = \sum_{i=1}^n S_i$ (this implies $S_n \sim \text{Exp}(q_{Y_{n-1}})$) and set $X_t = Y_n$ if $J_n \leq t < J_{n+1}$ and $X_t = \infty$ otherwise.

Construction 2:

- Let $(T_n^y)_{n \geq 1, y \in I}$ be iid Exp(1) random variables
- $Y_0 \sim \lambda$ and inductively define Y_n, S_n : if $Y_n = x$ then for $y \neq x$ define $S_{n+1}^y = \frac{T_{n+1}^y}{q_{xy}} \sim \text{Exp}(q_{xy})$ and $S_{n+1} = \inf_{y \neq x} S_{n+1}^y \sim \text{Exp}\left(\sum_{y \neq x} q_{xy}\right)$, and if $S_{n+1} = S_{n+1}^Z$ for some random Z (since the infimum is attained), take $Y_{n+1} = Z$ (if $q_x > 0$). If $q_x = 0$ take $Y_{n+1} = x$.

(Proof of equivalence: see Example Sheet)

Construction 3:

- For $x \neq y$, let $(N_t^{x,y})$ be independent Poisson Processes with rates q_{xy} respectively. Let $Y_0 \sim \lambda$, $J_0 = 0$ and define inductively:

$$J_{n+1} = \inf\{t > J_n : N_t^{Y_n, y} \neq N_{J_n}^{Y_n, y} \text{ for some } y \neq Y_n\},$$

$$Y_{n+1} = \begin{cases} y & \text{if } J_{n+1} < \infty \text{ and } N_{J_{n+1}}^{Y_n, y} \neq N_{J_n}^{Y_n, y} \\ x & \text{if } J_{n+1} = \infty \end{cases}.$$

For a birth process, we characterised when explosion happens. In general, the next theorem gives a sufficient condition:

Theorem 1.12. *Let X be Markov(λ, Q) on I . Then $\mathbb{P}(\zeta = \infty) = 1$ (non-explosive) if any of the following hold:*

- (a) I is finite;
- (b) $\sup_{x \in I} q_x < \infty$;
- (c) $X_0 = x$ and x is recurrent for the jump chain Y .

Proof. Note that (a) \Rightarrow (b) so it is enough to show in the cases we have (b) or (c). If (b) holds, set $q = \sup_{x \in I} q_x < \infty$. Since $S_n = \frac{T_n}{q_{X_{n-1}}}$, $S_n \geq \frac{T_n}{q}$. Hence

$$\zeta = \sum_{n=1}^{\infty} S_n > \frac{1}{q} \sum_{n=1}^{\infty} T_n = \infty \text{ almost surely (SLLN),}$$

i.e $\mathbb{P}(\zeta = \infty) = 1$.

Now suppose (c) holds. Let $(N_i)_{i \in I}$ be the times when the jump chain Y visits x . By the SLLN,

$$\zeta \geq \sum_{i=1}^{\infty} S_{N_i+1} = \sum_{i=1}^{\infty} \frac{T_{N_i+1}}{q_{N_i}} = \frac{1}{q_x} \sum_{i=1}^{\infty} T_{N_i+1} = \infty \text{ almost surely,}$$

i.e $\mathbb{P}(\zeta = \infty) = 1$. □

Example. Suppose $I = \mathbb{Z}$, $q_{i,i+1} = q_{i,i-1} = 2^{|i|}$ for all i . Then $p_{i,i+1} = p_{i,i-1} = 1/2$ and the jump chain is the symmetric simple random walk on \mathbb{Z} , which is recurrent. Hence X is non-explosive.

Example. Suppose $I = \mathbb{Z}$, $q_{i,i+1} = 2^{|i|+1}$, $q_{i,i-1} = 2^{|i|}$ so $q_i = 2^{|i|} + 2^{|i|+1}$. Then the jump chain Y is a simple random walk with $1/3$ probability of moving towards 0 and $2/3$ probability of moving away from 0, hence is transient. We have

$$\mathbb{E}[\zeta] = \mathbb{E} \left[\sum_{n=1}^{\infty} S_n \right] = \sum_{j \in \mathbb{Z}} \mathbb{E} \left[\sum_{k=1}^{V_j} S_{N_k^j+1} \right],$$

where V_j is the total number of visits to j and N_k^j is the time of the k th visit to j . Hence

$$\sum_{j \in \mathbb{Z}} \mathbb{E} \left[\sum_{k=1}^{V_j} S_{N_k^j+1} \right] = \sum_{j \in \mathbb{Z}} \mathbb{E}[V_j] \mathbb{E}[S_{N_1^j+1}] = \sum_{j \in \mathbb{Z}} \mathbb{E}[V_j] \frac{1}{q_j} = \sum_{j \in \mathbb{Z}} \frac{1}{3 \cdot 2^{|j|}} \mathbb{E}V_j.$$

Since $\mathbb{E}V_j \leq 1 + \mathbb{E}_j V_j = 1 + \mathbb{E}_0 V_0 := C < \infty$ (transience) we have

$$\sum_{j \in \mathbb{Z}} \frac{1}{3 \cdot 2^{|j|}} \mathbb{E}V_j \leq \sum_{j \in \mathbb{Z}} \frac{C}{2 \cdot 2^{|j|}} < \infty.$$

So $\mathbb{E}[\zeta] < \infty$ and $\mathbb{P}(\zeta < \infty) = 1$, i.e explosive.

Theorem 1.13 (Strong Markov Property). *Let X be Markov(λ, Q) and let T be a stopping time. Then conditional on $T < \zeta$ and $X_T = x$, the process $(X_{T+t})_{t \geq 0}$ is Markov(δ_x, Q) and independent of $(X_s)_{s \leq T}$.*

Proof. Omitted (uses measure theory, see Norris (6.5)). □

Kolmogorov's forward & backward equations

We work on a countable state space I .

Theorem 1.14. *Let X be a minimal right-continuous process with values in a countable set I . Let Q be a Q -matrix with jump matrix P . Then the following are equivalent:*

(a) X is a continuous-time Markov chain with generator Q .

(b) For all $n \geq 0$, $0 \leq t_0 \leq \dots \leq t_{n+1}$, and all states $x_0, \dots, x_{n+1} \in I$,

$$\mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n, \dots, X_{t_0} = x_0) = p_{x_n x_{n+1}}(t_{n+1} - t_n).$$

Where $(P(t)) = (p_{xy}(t))$ is the minimal non-negative solution to the backward equation

$$P'(t) = QP(t), \text{ with } P(0) = I.$$

(Minimality means that if \tilde{P} is another non-negative solution, we have $p_{xy}(t) \leq \tilde{p}_{xy}(t)$ for all t and all $x, y \in I$.) In fact, if the chain is non-explosive, the solution is unique.

(c) $P(t)$ is the minimal non-negative solution to the forward equation

$$P'(t) = P(t)Q, \text{ with } P(0) = I.$$

Note. We shall skip the proof of the equivalence of (c) (see Norris (2.8)).

Proof. First we show (a) \Rightarrow (b). If $(J_n)_{n \geq 1}$ denote the jump times, then

$$\mathbb{P}_x(X_t = y, J_1 > t) = \mathbb{1}(x = y)e^{-q_x t}.$$

Integrating over the values of $J_1 \leq t$ and using independence of the jump chain, for $z \neq x$,

$$\begin{aligned} \mathbb{P}_x(X_t = y, J_1 \leq t, X_{J_1} = z) &= \int_0^t q_x e^{-q_x s} \frac{q_{xz}}{q_x} p_{zy}(t-s) ds \\ &= \int_0^t e^{-q_x s} q_{xz} p_{zy}(t-s) ds \end{aligned}$$

Summing over all $z \neq x$ (and by the MCT),

$$\mathbb{P}_x(X_t = y, J_1 \leq t) = \int_0^t \sum_{z \neq x} e^{-q_x s} q_{xz} p_{zy}(t-s) ds.$$

So

$$p_{xy}(t) = \mathbb{P}_x(X_t = y) = e^{-q_x t} \mathbb{1}(x = y) + \int_0^t \sum_{z \neq x} e^{-q_x s} q_{xz} p_{zy}(t-s) ds.$$

And by a substitution

$$e^{q_x t} p_{xy}(t) = \mathbb{1}(x = y) + \int_0^t \sum_{z \neq x} e^{q_x u} q_{xz} p_{zy}(u) du.$$

Hence $p_{xy}(t)$ is a continuous function in t , and hence

$$\sum_{z \neq x} e^{q_x u} q_{xz} p_{zy}(u)$$

is a series of continuous functions, and is also uniformly convergence (Weierstrass-M test), so continuous. Hence $e^{q_x t} p_{xy}(t)$ is differentiable with derivative

$$e^{q_x t} (q_x p_{xy}(t) + p'_{xy}(t)) = \sum_{z \neq x} e^{q_x t} q_{xz} p_{zy}(t).$$

Thus

$$p'_{xy}(t) = \sum_z q_{xz} p_{zy}(t) \implies P'(t) = QP(t).$$

Now we show minimality: let \tilde{P} be another non-negative solution of the backward equation. We will show $p_{xy}(t) \leq \tilde{p}_{xy}(t)$ for all x, y, t . As before,

$$\begin{aligned} \mathbb{P}_x(X_t = y, t < J_{n+1}) &= \mathbb{P}_x(X_t = y, J_1 > t) + \mathbb{P}_x(X_t = y, J_1 \leq t < J_{n+1}) \\ &= e^{-q_x t} \mathbb{1}(x = y) + \sum_{z \neq x} \int_0^t q_x e^{-q_x s} \frac{q_{xz}}{q_x} \mathbb{P}_z(X_{t-s} = y, t-s < J_n) ds. \end{aligned}$$

Now, as \tilde{P} satisfies the backward equation, we get as before (retracing previous steps)

$$\tilde{p}_{xy}(t) = e^{-q_x t} \mathbb{1}(x = y) + \sum_{z \neq x} \int_0^t e^{-q_x s} q_{xz} \tilde{p}_{zy}(t-s) ds. \quad (*)$$

Now we prove by induction that

$$\mathbb{P}_x(X_t = y, t < J_n) \leq \tilde{p}_{xy}(t) \text{ for all } n.$$

For $n = 1$,

$$e^{-q_x t} \mathbb{1}(x = y) \leq \tilde{p}_{xy}(t) \text{ by } (*).$$

Assume true for some $n \in \mathbb{N}$. Then for $n + 1$,

$$\mathbb{P}_x(X_t = y, t < J_{n+1}) \leq e^{-q_x t} \mathbb{1}(x = y) + \sum_{z \neq x} \int_0^t q_{xz} e^{-q_x s} \tilde{p}_{zy}(t-s) ds = \tilde{p}_{xy}(t).$$

So it holds for all n . Hence

$$\lim_{n \rightarrow \infty} \mathbb{P}_x(X_t = y, t < J_n) = \mathbb{P}_x(X_t = y, t < \zeta) \leq \tilde{p}_{xy}.$$

(Since $J_n \uparrow \zeta$.) Now by minimality,

$$p_{xy}(t) = \mathbb{P}_x(X_t = y) = \mathbb{P}_x(X_t = y, t < \zeta) \leq \tilde{p}_{xy}(t).$$

□

Finite state space:

Definition. If A is a finite-dimensional square matrix, its matrix exponential is given by

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2!} + \dots$$

Claim. For any $r \times r$ matrix A , the exponential e^A is an $r \times r$ matrix. If A_1 and A_2 commute, then $e^{A_1+A_2} = e^{A_1}e^{A_2}$.

Proof. Example Sheet. □

Proposition 1.15. Let Q be a Q -matrix on a finite set I and $P(t) = e^{tQ}$. Then

- (i) $P(t+s) = P(t)P(s)$ for all s, t ;
- (ii) $(P(t))_{t \geq 0}$ is the unique solution to the forward equation $P'(t) = P(t)Q$, $P(0) = I$;
- (iii) $(P(t))_{t \geq 0}$ is the unique solution to the backward equation $P'(t) = QP(t)$, $P(0) = I$;
- (iv) For $k = 0, 1, 2, \dots$, $\left(\frac{d}{dt}\right)^k P(t) \Big|_{t=0} = Q^k$.

Proof.

- (i) Since tQ and sQ commute, $\exp((t+s)Q) = \exp(tQ)\exp(sQ)$.
- (ii) The sum in e^{tQ} has infinite radius of convergence, hence we can differentiate term by term.
- (iii) Same as (ii).
- (iv) Same again.

Now we'll show uniqueness in (ii) and (iii). If \tilde{P} is another solution to the forward equation, $\tilde{P}'(t) = \tilde{P}(t)Q$, $\tilde{P}(0) = I$, then

$$\begin{aligned}\frac{d}{dt} \left(\tilde{P}(t)e^{-tQ} \right) &= \tilde{P}'(t)e^{-tQ} + \tilde{P}(t) (-Qe^{-tQ}) \\ &= \tilde{P}(t)Qe^{-tQ} - \tilde{P}(t)Qe^{-tQ} = 0\end{aligned}$$

So $\tilde{P}(t)e^{-tQ}$ is a constant matrix. Since $\tilde{P}(0) = I$, this implies $\tilde{P}(t) = e^{tQ}$. The same thing works for the backward equation. \square

Example. Let $Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}$. To find $p_{11}(t)$, we can diagonalise $Q = PDP^{-1}$ for a diagonal matrix

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

so

$$e^{tQ} = Pe^{tD}P^{-1} = P \begin{pmatrix} e^{t\lambda_1} & 0 & 0 \\ 0 & e^{t\lambda_2} & 0 \\ 0 & 0 & e^{t\lambda_3} \end{pmatrix} P^{-1}.$$

i.e $p_{11}(t) = ae^{t\lambda_1} + be^{t\lambda_2} + ce^{t\lambda_3}$, which we can solve by considering $p_{11}(0), p'_{11}(0), p''_{11}(0)$.

Theorem 1.16. *Let I be a finite state space and Q be a matrix. Then it is a Q -matrix iff $P(t) = e^{tQ}$ is a stochastic matrix for all t .*

Proof. For t sufficiently small, $p(t) = e^{tQ} = I + tQ + \mathcal{O}(t^2)$, so for all $x \neq y$, $q_{xy} \geq 0$ iff $p_{xy}(t) \geq 0$ for all t sufficiently small.

Since $P(t) = (P(t/n))^n$ for all n , we get $q_{xy} \geq 0$ for all $x \neq y$ iff $p_{xy}(t) \geq 0$ for all $t \geq 0$.

Assume now that Q is a Q -matrix, i.e. $\sum_y q_{xy} = 0$ for all x . Then $\sum_y (Q^n)_{xy} = \sum_y \sum_z (Q^{n-1})_{xz} Q_{zy} = \sum_z Q_{xz}^{n-1} \sum_y Q_{zy} = 0$. Hence $Q^n \mathbf{1} = Q^{n-1} Q \mathbf{1} = 0$ ($\mathbf{1}$ is vector with all entries 1). Hence, since

$$p_{xy}(t) = \delta_{xy} + \sum_{k=1}^{\infty} \frac{t^k}{k!} (Q^k)_{xy}$$

we have $\sum_y p_{xy}(t) = 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_y (Q^k)_{xy} = 1$. i.e $P(t)$ is a stochastic matrix.

Assume now that $P(t)$ is a stochastic matrix. Then as $Q = \left. \frac{d}{dt} \right|_{t=0} P(t)$, we have

$$\sum_y q_{xy} = \left. \frac{d}{dt} \right|_{t=0} \sum_y p_{xy}(t) = 0.$$

i.e Q is a Q -matrix. □

Theorem 1.17. *Let X be a right-continuous process with values in a finite set I , and let Q be a Q -matrix on I . Then the following are equivalent*

- (a) *The process X is Markov with generator Q (Markov(Q));*
- (b) *(infinitesimal definition) Conditional on $X_s = x$, the process $(X_{s+t})_{t \geq 0}$ is independent of $(X_r)_{r \leq s}$ and uniformly in t as $h \downarrow 0$, for all x, y*

$$\mathbb{P}(X_{t+h} = y | X_t = x) = \mathbb{1}(x = y) + q_{xy}h + o(h)$$

- (c) *For all $n \geq 0$, $0 \leq t_0 \leq \dots \leq t_n$ and all states x_0, \dots, x_n ,*

$$\mathbb{P}(X_{t_n} = x_n | X_{t_0} = x_0, \dots, X_{t_{n-1}} = x_{n-1}) = p_{x_{n-1}, x_n}(t_n - t_{n-1})$$

where $(p_{xy}(t))$ is the solution to the forward equation $P'(t) = P(t)Q$, $P(0) = I$.

Proof. We have already shown (a) \iff (b) (from countable setting), so it is enough to show (b) \iff (c).

First we show (c) \implies (b). $P(t) = e^{tQ}$ is the solution (as I is finite). As $t \downarrow 0$, $P(t) = I + tQ + \mathcal{O}(t^2)$. Thus for all $t > 0$ and as $h \downarrow 0$, $\forall x, y$,

$$\mathbb{P}(X_{t+h} = y | X_t = x) = \mathbb{P}(X_h = y | X_0 = x) = p_{xy}(h) = \delta_{xy} + hq_{xy} + o(h).$$

Now we show (b) \Rightarrow (c). We have

$$p_{xy}(t+h) = \sum_z p_{xz}(t)(\mathbb{1}(z=y) + q_{zy}h + o(h)).$$

So

$$\frac{p_{xy}(t+h) - p_{xy}(t)}{h} = \sum_z p_{xz}(t)q_{zy} + o(1).$$

As $h \downarrow 0$,

$$p'_{xy}(t) = \sum_z p_{xz}(t)q_{zy} = (P(t)Q)_{xy}.$$

□

Remark. To get the backward equation we could write

$$p_{xy}(t+h) = \sum_z p_{xz}(h)p_{zy}(t)$$

and continue similarly.

2 Qualitative Properties of Continuous Time Markov Chains

We have minimal chains, and countable state space.

Class Structure

Definition. For states $x, y \in I$, write $x \rightarrow y$ (“ x leads to y ”) if $\mathbb{P}_x(X_t = y \text{ for some } t \geq 0) > 0$. We write $x \leftrightarrow y$ (“ x communicates with y ”) if $x \rightarrow y$ and $y \rightarrow x$. Clearly this is an equivalence relation and we call the equivalence classes *communicating classes*. We define *irreducibility*, *closed class* and *absorbing states* exactly as in discrete Markov Chains.

Proposition 2.1. Let X be Markov(Q) with transition semigroup $(P(t))_{t \geq 0}$. For any 2 states $x, y \in I$, the following are equivalent

- (a) $x \rightarrow y$;
- (b) $x \rightarrow y$ for the jump chain;
- (c) $q_{x_0 x_1} \cdots q_{x_{n-1} x_n} > 0$ for some $x = x_0, x_1, \dots, x_{n-1}, x_n = y$;
- (d) $p_{xy}(t) > 0$ for all $t > 0$;
- (e) $p_{xy}(t) > 0$ for some $t > 0$.

Proof. Clearly (d) \Rightarrow (e) \Rightarrow (a) \Rightarrow (b). Now we show (b) \Rightarrow (c). Since $x \rightarrow y$ for the jump chain, there exist $x_0 = x, x_1, \dots, x_{n-1}, x_n = y \in I$ such that

$$p_{x_0 x_1} p_{x_1 x_2} \dots p_{x_{n-1} x_n} > 0.$$

Hence $q_{x_0 x_1} q_{x_1 x_2} \dots q_{x_{n-1} x_n}$ since $q_{xy}/q_x = p_{xy}$.

Now we show (c) \Rightarrow (d). For any 2 states w, z with $q_{wz} > 0$, and for any $t > 0$,

$$p_{wz}(t) \geq \mathbb{P}_w(J_1 \leq t, Y_1 = z, S_2 > t) = (1 - e^{-q_w t}) \frac{q_{wz}}{q_w} e^{-q_z t} > 0.$$

i.e $q_{wz} > 0$ implies $q_{wz}(t) > 0$ for all t . Hence if (c) holds, $p_{x_i x_{i+1}}(t) > 0$ for all t and all $0 \leq i \leq n-1$. Then $p_{xy}(t) = p_{x_0 x_1}(t/n) p_{x_1 x_2}(t/n) \dots p_{x_{n-1} x_n}(t/n) > 0$. \square

Hitting times

Definition. Let Y be the jump chain associated with X , and $A \subseteq I$. Set $T_A = \inf\{t > 0 : X_t \in A\}$, $H_A = \inf\{n \geq 0 : Y_n \in A\}$, $h_A(x) = \mathbb{P}_x(T_A < \infty)$ (hitting probability), $k_A(x) = \mathbb{E}_x T_A$ (mean hitting time).

Note. The hitting probability for X is the same as that for Y but the mean hitting times will differ in general.

Theorem 2.2. $(h_A(x))_{x \in I}$ and $(k_A(x))_{x \in I}$ are the minimal non-negative solutions to

$$\begin{cases} h_A(x) = 1 & \forall x \in A \\ Qh_A(x) = \sum_y q_{xy} h_A(y) = 0 & \forall x \notin A \end{cases}$$

and

$$\begin{cases} k_A(x) = 0 & \forall x \in A \\ Qk_A(x) = \sum_y q_{xy} k_A(y) = -1 & \forall x \notin A \end{cases}$$

respectively (assume $q_x > 0$ for all $x \notin A$).

Proof. The hitting probabilities are the same as those for the jump chain. Hence $h_A(x) = 1$ for all $x \in A$ and $h_A(x) = \sum_{y \neq x} p_{xy} h_A(y)$ for all $x \notin A$. Hence for all $x \notin A$

$$q_x h_A(x) = \sum_{y \neq x} h_A(y) q_{xy} \implies \sum_y h_A(y) q_{xy} = 0.$$

Clearly if $x \in A$, $T_A = 0$, so $k_A(x) = 0$. Let $x \notin A$. Then $J_1 \leq T_A$, and hence

$$\begin{aligned} k_A(x) &= \mathbb{E}_x T_A \\ &= \mathbb{E}_x J_1 + \mathbb{E}_x (T_A - J_1) \\ &= \mathbb{E}_x J_1 + \sum_{y \neq x} \mathbb{E}_x (T_A - J_1 | Y_1 = y) p_{xy} \\ &= \frac{1}{q_x} + \sum_{y \neq x} k_A(y) \frac{q_{xy}}{q_x}. \end{aligned}$$

Therefore

$$q_x k_A(x) = 1 + \sum_{y \neq x} q_{xy} k_A(y) \implies \sum_y q_{xy} k_A(y) = -1.$$

The minimality of solutions is as in the discrete chain. □

Recurrence and Transience

Definition. The state x is called *recurrent* for X if

$$\mathbb{P}(\{t : X_t = x\} \text{ is unbounded}) = 1.$$

The state x is called *transient* if

$$\mathbb{P}(\{t : X_t = x\} \text{ is unbounded}) = 0.$$

Remark. If X explodes with positive probability starting from x , i.e. $\mathbb{P}(\zeta < \infty) > 0$, then $\sup_t \{t : X_t = x\} \leq \zeta < \infty$ with positive probability so x cannot be recurrent.

Theorem 2.3. Let X be Markov(Q) with jump chain Y . Then

- (a) If x is recurrent for Y , then x is recurrent for X ;
- (b) If x is transient for Y , then x is transient for X ;
- (c) Every state is either recurrent or transient;
- (d) Recurrence and transience are class properties.

Proof. (a) & (b) will imply (c) & (d) through the results for the discrete chain. So we prove (a) and (b).

First we prove (a). Suppose x is recurrent for Y and $X_0 = x$. Then X is not explosive, i.e $\mathbb{P}(\zeta = \infty) = 1$, so $J_n \rightarrow \infty$ with probability 1 (starting from x). Since $X_{J_n} = Y_n$ for all n , and Y visits x infinitely often with probability 1, $\{t : X_t = x\}$ is unbounded with probability 1.

Now we prove (b). If x is transient for Y , $q_x > 0$ (otherwise x is an absorbing state). Also, almost surely there is a last visit to x for Y , i.e

$$N := \sup\{n : Y_n = x\} < \infty \text{ almost surely.}$$

Also, $J_{N+1} < \infty$ almost surely (as $q_x > 0$) and if $t \in \{s : X_s = x\}$, then $t \leq J_{N+1}$, i.e $\sup\{s : X_s = x\} \leq J_{N+1} < \infty$ almost surely. \square

Like in the discrete-time chain, $\sum_{n \geq 1} p_{xx}(n) = \infty$ implies x is recurrent; and $\sum_{n \geq 1} p_{xx}(n) < \infty$ implies x is transient.

Theorem 2.4. x is recurrent for X if and only if $\int_0^\infty p_{xx}(t)dt = \infty$, and x is transient for X if and only if $\int_0^\infty p_{xx}(t)dt < \infty$.

Proof. If $q_{xx} = 0$, then x is absorbing, i.e $p_{xx}(t) = 1$ for all t and $\int_0^\infty p_{xx}(t)dt = \infty$. Assume $q_x > 0$. Then

$$\begin{aligned} \int_0^\infty p_{xx}(t)dt &= \int_0^\infty \mathbb{E}[\mathbb{1}(X_t = x)]dt \\ &= \mathbb{E}_x \left[\int_0^\infty \mathbb{1}(X_t = x)dt \right] && \text{(Fubini)} \\ &= \mathbb{E}_x \left[\sum_{n=0}^\infty \mathbb{1}(Y_n = x) S_{n+1} \right] \\ &= \sum_{n=0}^\infty \mathbb{E}_x [\mathbb{1}(Y_n = x) S_{n+1}] && \text{(Fubini)} \\ &= \sum_{n=0}^\infty \mathbb{P}_x(Y_n = x) \mathbb{E}_x[S_{n+1} | Y_n = x] \\ &= \sum_{n=0}^\infty p_{xx}(n) \frac{1}{q_x}. \end{aligned}$$

\square

Invariant Distributions

Definition. For a discrete Markov Chain Y , π is an *invariant measure* for Y if $\pi P = \pi$. If in addition $\sum \pi_i = 1$, π is called a *invariant distribution*. Then if $Y_0 \sim \pi$, $Y_n \sim \pi$ for all $n \geq 1$.

Recall:

Theorem 2.5. If Y is a discrete time Markov Chain which is irreducible, recurrent and $x \in I$. Then

$$\nu^x(y) = \mathbb{E}_x \left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y) \right] \text{ where } H_x = \inf\{n \geq 1 : Y_n = x\}.$$

Then $\nu^x(\cdot)$ is an invariant measure and $0 < \nu^x(y) \leq 1$ for all y , $\nu^x(x) = 1$.

Theorem 2.6. If Y is irreducible, λ is any invariant measure with $\lambda(x) = 1$, then

$$\lambda(y) \geq \nu^x(y) \text{ for all } y.$$

If Y is recurrent then $\lambda(y) = \nu^x(y)$ for all y .

Definition. Let $X \sim \text{Markov}(Q)$ and let λ be a measure. Then λ is called invariant/infinitesimally invariant if $\lambda Q = 0$.

Lemma 2.7. If $|I|$ is finite, then $\lambda Q = 0$ if and only if $\lambda P(s) = \lambda$ for all $s \geq 0$.

Proof. $P(s) = e^{sQ}$ since I is finite. If $\lambda Q = 0$, then

$$\lambda P(s) = \lambda e^{sQ} = \lambda \sum_{k=0}^{\infty} \frac{(sQ)^k}{k!} = I.$$

If $\lambda P(s) = \lambda$ for all s , then

$$\lambda Q = \lambda P'(0) = \left. \frac{d}{ds} (\lambda P(s)) \right|_{s=0} = \left. \frac{d}{ds} \lambda \right|_{s=0} = 0.$$

□

Lemma 2.8. Let X be $\text{Markov}(Q)$ and Y its jump chain. π is invariant for X if and only if μ defined by $\mu_x = q_x \pi_x$ is invariant for Y (i.e. $\pi Q = 0$ if and only if $\mu P = \mu$).

Proof. Since $q_x (p_{xy} - \delta_{xy}) = q_{xy}$,

$$\begin{aligned} (\pi Q)_y &= \sum_{x \in I} \pi_x q_{xy} = \sum_{x \in I} \pi_x q_x (p_{xy} - \delta_{xy}) \\ &= \sum_{x \in I} \mu_x (p_{xy} - \delta_{xy}) \\ &= \sum_x \mu_x p_{xy} - \mu_y \\ &= (\mu P)_y - \mu_y. \end{aligned}$$

□

Theorem 2.9. Let X be irreducible & recurrent, with generator Q . Then X has an invariant measure, which is unique up to scalar multiplication.

Proof. Assume $|I| > 1$. Then by irreducibility, $q_x > 0$ for all x . For Y , $\nu^x(y) = \mathbb{E}_x \left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y) \right]$ where $H_x = \inf\{n \geq 1 : Y_n = x\}$ is an invariant measure as Y is irreducible & recurrent (since X is), hence ν^x is an invariant measure for Y which is unique up to scalar multiplication. By the previous lemma, $\frac{\nu^x(y)}{q_y}$ is an invariant measure for X , and also unique up to scalar multiplication. □

Definition. Let $T_x = \inf\{t \geq J_1 : X_t = x\}$ be the first return time to x .

Lemma 2.10. Assume $q_y > 0$. Define

$$\mu^x(y) = \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y) dt \right].$$

Then $\mu^x(y) = \frac{\nu^x(y)}{q_y}$.

Proof.

$$\begin{aligned}
 \mu^x(y) &= \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y) dt \right] \\
 &= \mathbb{E}_x \left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y) S_{n+1} \right] \\
 &= \mathbb{E}_x \left[\sum_{n=0}^{\infty} S_{n+1} \mathbb{1}(Y_n = y, n \leq H_x - 1) \right] \\
 &= \sum_{n=0}^{\infty} \mathbb{E}_x [S_{n+1} | Y_n = y, n \leq H_x - 1] \mathbb{P}_x(Y_n = y, n \leq H_x - 1)
 \end{aligned}$$

Since $\{n < H_x\}^c = \{H_x \leq n\} \in \sigma\{Y_1, \dots, Y_n\}$ (i.e depends on Y_1, \dots, Y_n only) its a stopping time so the Strong Markov Property says

$$\begin{aligned}
 \mu^x(y) &= \sum_{n=0}^{\infty} \mathbb{E}_x [S_{n+1} | Y_n = y] \mathbb{P}_x(Y_n = y, n \leq H_x - 1) \\
 &= \sum_{n=0}^{\infty} \frac{1}{q_y} \mathbb{P}_x(Y_n = y, n \leq H_x - 1) \\
 &= \frac{1}{q_y} \sum_{n=0}^{\infty} \mathbb{E}_x [\mathbb{1}(Y_n = y, n \leq H_x - 1)] \\
 &= \frac{1}{q_y} \mathbb{E} \left[\sum_{n=0}^{\infty} \mathbb{1}(Y_n = y, n \leq H_x - 1) \right] \\
 &= \frac{1}{q_y} \mathbb{E}_x \left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y) \right] \\
 &= \frac{\nu^x(y)}{q_y}.
 \end{aligned}$$

□

Definition. A recurrent state x is called *positive recurrent* if

$$m_x = \mathbb{E}_x T_x < \infty.$$

Otherwise, we call x *null recurrent*.

Theorem 2.11. Let $X \sim \text{Markov}(Q)$ be irreducible. Then the following are equivalent

- (a) Every state is positive recurrent;
- (b) Some state is positive recurrent;

(c) X is non-explosive and has an invariant distribution.

Also, when (c) holds, the invariant distribution λ is given by $\lambda(x) = \frac{1}{q_x m_x}$ for all x .

Proof. Clearly (a) \Rightarrow (b). Now we show (b) \Rightarrow (c). Assume without loss of generality that $q_x > 0$. Let x be a positive recurrent state. Then all states are recurrent, so Y is recurrent and the chain is non-explosive starting from any y . As Y is recurrent, ν^x is an invariant measure for Y . So $\mu^x = \frac{\nu^x}{q_y}$ (as defined previously) is an invariant measure for X . Also

$$\mu_x(y) = \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y) dt \right],$$

so

$$\begin{aligned} \sum_{y \in I} \mu^x(y) &= \mathbb{E}_x \left[\int_0^{T_x} \sum_{y \in I} \mathbb{1}(X_t = y) dt \right] \\ &= \mathbb{E}_x T_x < \infty. \end{aligned}$$

So μ_x is normalisable, and $\frac{\mu_x}{\mathbb{E}_x T_x}$ is an invariant distribution for X .

Now we show (c) \Rightarrow (a). By a previous lemma, the measure $\beta(y) = \lambda(y)q_y$ is an invariant measure for Y . Since $\sum_{y \in I} \lambda(y) = 1$, $\lambda(x) > 0$ for some x . Since Y is irreducible, for any $y \in I$, $x \rightarrow y$, i.e. $p_{xy}(n) > 0$ for some n . As β is invariant for Y , $\beta P^n = \beta$. So

$$\lambda(y)q_y = \beta y = \sum_{z \in I} \beta_z p_{zy}(n) \geq \beta_x p_{xy}(n) = \lambda(x)q_x p_{xy}(n) > 0$$

so $\lambda(y) > 0$ for all y . Fix some $x \in I$. Then $\lambda(x) > 0$ so define $a^x(y) = \frac{\beta(y)}{\lambda(x)q_x}$ for all $y \in I$, which is invariant for Y as a scalar multiple of $\beta(y)$, and $a^x(x) = 1$. By the theorem for discrete-time chains $a^x(y) \geq \nu^x(y)$ for all $y \in I$, where $\nu^x(y) = \mathbb{E}_x \left[\sum_{n=0}^{H_x-1} \mathbb{1}(Y_n = y) \right]$ and where $H_x = \inf\{n \geq 1 : Y_n = x\}$.

Also if $\mu^x(y) = \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y) dt \right]$ then $\mu^x(y) = \frac{\nu^x(y)}{q_y}$ and

$$\begin{aligned} \sum_{y \in I} \mu^x(y) &= \mathbb{E}_x \left[\int_0^{T_x} \sum_{y \in I} \mathbb{1}(X_t = y) dt \right] \\ &= \mathbb{E}_x T_x = m_x \quad (\text{as } X \text{ is non-explosive}) \end{aligned}$$

Then

$$\begin{aligned}
 \mu_x &= \sum_y \mu^x(y) = \sum_y \frac{\nu^x(y)}{q_y} \leq \sum_y \frac{a^x(y)}{q_y} \\
 &= \sum_y \frac{\beta(y)}{\lambda(x)q_xq_y} \\
 &= \sum_y \frac{\lambda(y)q_y}{\lambda(x)q_xq_y} \\
 &= \frac{1}{\lambda(x)q_x} \sum_y \lambda(y) \\
 &= \frac{1}{\lambda(x)q_x} < \infty.
 \end{aligned}$$

Hence x is positive recurrent. As x was arbitrary this means all states are positive recurrent.

Also, if (c) holds, then X is recurrent, so Y is recurrent. Hence $a^x(y) = v^x(y)$ for all y . Therefore $m_x = \frac{1}{\lambda(x)q_x}$ as the previous inequality becomes equality. \square

Example. On \mathbb{Z}^+ , suppose $q_{i,i+1} = \lambda q_i$, $q_{i,i-1} = \mu q_i$ and $q_{ii} = -(\lambda + \mu)q_i$ and $q_{i,j} = 0$ for all other j (an example of a Birth & Death process). We have transition probabilities $p_{i,i+1} = \frac{\lambda}{\lambda + \mu}$ and $p_{i,i-1} = \frac{\mu}{\lambda + \mu}$. Then $(\lambda/\mu)^i$ is an invariant measure for Y . Then $\pi_i = \frac{1}{q_i}(\lambda/\mu)^i$ is invariant for X . So if $q_i = 2^i$ and $\lambda = \frac{3\mu}{2}$, then $\pi_i = (3/4)^i$ is invariant for X . Also $\sum_{i=0}^{\infty} \pi_x < \infty$ so X has an invariant distribution. Since $\lambda > \mu$, the chain is transient for Y and so is transient for X . If X were non-explosive then by the previous theorem it would be positive recurrent, hence X must be explosive.

Lemma 2.12. *Let X be a continuous-time Markov chain. Fix $t > 0$ and set $Z_n = X_{nt}$. Then $(Z_n)_{n=0}^{\infty}$ is a discrete-time Markov chain. Then x is recurrent for X if and only if x is recurrent for Z .*

Proof. Example Sheet. □

Theorem 2.13. *Let $X \sim \text{Markov}(Q)$ be recurrent, irreducible and λ be a measure. Then $\lambda Q = 0$ if and only if $\lambda P(s) = \lambda$ for all $s > 0$.*

Proof. Any measure λ such that $\lambda Q = 0$ is unique up to scalar multiplication (by a theorem proved previously).

Any measure λ such that $\lambda P(s) = \lambda$ for all s is unique up to scalar multiplication. Indeed, fix $s = 1$ so $\lambda P(1) = \lambda$. Then $(X_n)_{n=0}^{\infty}$ is a discrete time chain with transition matrix $P(1)$, and is irreducible, recurrent by the previous lemma. It also has λ as an invariant measure, hence unique (up to scalar multiplication).

So it is enough to show $\mu^x Q = 0$ and $\mu^x P(s) = \mu^x$ for all s where $\mu^x(y) = \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y) dt \right]$.

Also $\mu^x(y) = \frac{\nu^x(y)}{q_y}$ and since X is recurrent, Y is recurrent so ν^x is an invariant measure for Y . So μ^x is an invariant measure for X , i.e $\mu^x Q = 0$.

Also, by the Strong Markov Property,

$$\mathbb{E}_x \left[\int_0^s \mathbb{1}(X_t = y) dt \right] = \mathbb{E}_x \left[\int_{T_x}^{T_x+s} \mathbb{1}(X_t = y) dt \right]. \quad (*)$$

Thus

$$\begin{aligned}
\mu^x(y) &= \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_t = y) dt \right] \\
&= \mathbb{E}_x \left[\int_0^s \mathbb{1}(X_t = y) dt \right] + \mathbb{E}_x \left[\int_s^{T_x} \mathbb{1}(X_t = y) dt \right] \\
&= \mathbb{E}_x \left[\int_{T_x}^{T_x+s} \mathbb{1}(X_t = y) dt \right] + \mathbb{E}_x \left[\int_s^{T_x} \mathbb{1}(X_t = y) dt \right] \\
&= \mathbb{E}_x \left[\int_s^{T_x+s} \mathbb{1}(X_t = y) dt \right] \\
&= \mathbb{E}_x \left[\int_0^\infty \mathbb{1}(X_{u+s} = y, u < T_x) du \right] \quad (\text{letting } t = u + s) \\
&= \int_0^\infty \mathbb{P}_x(X_{u+s} = y, u < T_x) du \\
&= \int_0^\infty \sum_{z \in I} \mathbb{P}_x(X_u = z, X_{u+s} = y, u < T_x) du \\
&= \sum_{z \in I} p_{zy}(s) \mathbb{E}_x \left[\int_0^{T_x} \mathbb{1}(X_u = z) dy \right] \\
&= \sum_{z \in I} \mu^x(z) p_{zy}.
\end{aligned}$$

i.e $\mu^x = \mu^x P(s)$. Since s was arbitrary, $\mu^x = \mu^x P(s)$ for all s . \square

Convergence to Equilibrium

Lemma 2.14. *For the semigroup $P(t)$ and all $t \geq 0$, $h \geq 0$,*

$$|p_{xy}(t+h) - p_{xy}(t)| \leq 1 - e^{-q_x h} \leq q_x h.$$

Proof.

$$\begin{aligned}
 |p_{xy}(t+h) - p_{xy}(t)| &= \left| \sum_z p_{xz}(h)p_{zy}(t) - p_{xy}(t) \right| \\
 &= \left| \underbrace{\sum_{z \neq x} p_{xz}(h)p_{zy}(t)}_{\in [0, 1-p_{xx}(h)]} - \underbrace{p_{xy}(t)(1-p_{xx}(h))}_{\in [0, 1-p_{xx}(h)]} \right| \\
 &\leq 1 - p_{xx}(h) \\
 &= \mathbb{P}_x(X(h) \neq x) \\
 &\leq \mathbb{P}_x(J_1 \leq h) \\
 &= 1 - e^{-q_x h}
 \end{aligned}$$

□

Theorem 2.15. *Let $X \sim \text{Markov}(Q)$ be irreducible, non-explosive, and let λ be an invariant distribution. Then for all $x, y \in I$, $p_{xy}(t) \rightarrow \lambda(y)$ as $t \rightarrow \infty$.*

Proof. Fix $\varepsilon > 0$. Fix $h > 0$ such that $q_x h < \varepsilon/2$. Consider the discrete time Markov Chain $(Z_n) = (X_{nh})_{n \geq 0}$. Then (Z_n) is irreducible and aperiodic ($p_{xy}(h) > 0$ for all x, y by irreducibility). As X is positive recurrent (non-explosive and has invariant distribution), $\lambda P(h) = \lambda$, so λ is an invariant distribution for Z_n .

By a discrete-time Markov Chain result, for all x, y , $p_{xy}(nh) \rightarrow \lambda(y)$ as $n \rightarrow \infty$. Hence there exists n_0 such that for all $n \geq n_0$, $|p_{xy}(nh) - \lambda(y)| < \varepsilon/2$. Let $t \geq n_0 h$. Then there exists $n \geq n_0$ such that $nh \leq t < (n+1)h$. So

$$|p_{xy}(t) - p_{xy}(nh)| \leq q_x(t - nh) \leq q_x h < \varepsilon/2.$$

Thus for all $n \geq n_0 h$,

$$|p_{xy}(t) - \lambda(y)| \leq |p_{xy}(t) - p_{xy}(nh)| + |p_{xy}(nh) - \lambda(y)| < \varepsilon.$$

□

Ergodic Theory

Theorem 2.16. *Let $X \sim \text{Markov}(\lambda, Q)$ be irreducible. Then*

$$\frac{1}{t} \int_0^t \mathbb{1}(X_s = x) ds \rightarrow \frac{1}{q_x m_x} \text{ as } t \rightarrow \infty \text{ almost surely.}$$

If X is positive recurrent \mathcal{E} π is the unique invariant distribution and $f : I \rightarrow \mathbb{R}$ is bounded, then

$$\frac{1}{t} \int_0^t f(X_s) ds \rightarrow \sum_{x \in I} f(x) \pi(x)$$

Proof. Not given. □

Note. The second limit can be justified by

$$\begin{aligned} \frac{1}{t} \int_0^t f(X_s) ds &= \frac{1}{t} \int_0^t \sum_{x \in I} f(x) \mathbb{1}(X_s = x) ds \\ &= \sum_{x \in I} f(x) \left(\frac{1}{t} \int_0^t \mathbb{1}(X_s = x) ds \right) \\ &\rightarrow \sum_{x \in I} f(x) \pi(x). \end{aligned}$$

Reversibility

Theorem 2.17. *Let $X \sim \text{Markov}(Q)$ be irreducible and non-explosive with invariant distribution π . Let $X_0 \sim \pi$. Fix $T > 0$ and set $\hat{X}_t = X_{T-t}$ for $0 \leq t \leq T$. Then $\hat{X} \sim \text{Markov}(\hat{Q})$ and has invariant distribution π where $\hat{q}_{xy} = \pi(y) \frac{q_{yx}}{\pi(x)}$. Also \hat{Q} is irreducible and non-explosive (i.e $Z \sim \text{Markov}(\hat{Q})$ is non-explosive).*

Proof. Note that \hat{Q} is indeed a Q -matrix: $\hat{q}_{xy} \geq 0$ for all x, y and $\sum_y \hat{q}_{xy} = \frac{1}{\pi(x)} \sum_y \pi(y) q_{yx} = \frac{1}{\pi(x)} (\pi Q)_x = 0$. Also \hat{Q} is irreducible (as Q is). Also $(\pi \hat{Q})_y = \sum_x \pi(x) \hat{q}_{xy} = \sum_x \pi(y) q_{yx} = 0$, so π is invariant for \hat{Q} .

Now, let $0 = t_0 \leq t_1 \leq \dots \leq t_n = T$, $x_1, \dots, x_n \in I$, let $s_i = t_i - t_{i-1}$. Then

$$\begin{aligned} \mathbb{P}(\hat{X}_{t_0} = x_0, \dots, \hat{X}_{t_n} = x_n) &= \mathbb{P}(X_0 = x_n, \dots, X_{T-t_1} = x_1, X_T = x_0) \\ &= \pi(x_n) p_{x_n x_{n-1}}(s_n) \dots p_{x_1 x_0}(s_1). \end{aligned}$$

Define $\hat{p}_{xy}(t) = \frac{\pi(y)}{\pi(x)} p_{yx}(t)$ so

$$\begin{aligned} \pi(x_n) p_{x_n x_{n-1}}(s_n) \dots p_{x_1 x_0}(s_1) &= \pi(x_n) \hat{p}_{x_{n-1} x_n}(s_n) \frac{\pi(x_{n-1})}{\pi(x_n)} \dots \hat{p}_{x_0 x_1}(s_1) \frac{\pi(x_0)}{\pi(x_1)} \\ &= \pi(x_0) \hat{p}_{x_0 x_1}(s_1) \dots \hat{p}_{x_{n-1} x_n}(s_n). \end{aligned}$$

So \hat{X} is Markov with transition semigroup $(\hat{P}(t))_{t \geq 0}$. Need to show that $\hat{P}(t)$ is the minimal non-negative solution to the Kolmogorov backward equation with \hat{Q} , that is $(\hat{P}(t))' = \hat{Q} \hat{P}(t)$.

Indeed,

$$\begin{aligned} \hat{p}'_{xy}(t) &= \frac{\pi(x)}{\pi(y)} p'_{yx}(t) \\ &= \frac{\pi(y)}{\pi(x)} \sum_z p_{yz}(t) q_{zx} && (\text{Kolmogorov forward eq for } P) \\ &= \frac{\pi(y)}{\pi(x)} \sum_z \frac{\pi(z)}{\pi(y)} \hat{p}_{zy}(t) q_{yx} \\ &= \frac{1}{\pi(x)} \sum_z \pi(x) \hat{q}_{xz} \hat{p}_{zy}(t) \\ &= (\hat{Q} \hat{P})_{xy}. \end{aligned}$$

Suppose R is another solution to the Kolmogorov forward equation: $R'(t) = \hat{Q} R(t)$. Then defining $\bar{R}_{xy}(t) = \frac{\pi(y)}{\pi(x)} R_{yx}(t)$ then as before \bar{R} satisfies $\bar{R}'(t) = \bar{R}(t) Q$. But we know that P is the minimal solution to this, so \hat{P} is minimal for the forward equation.

Now we show \hat{Q} is non-explosive. Indeed, X is irreducible and non-explosive with invariant distribution π , so X is (positive) recurrent. Hence $\pi P(t) = \pi$ for all t . Thus

$$\sum_y \hat{p}_{xy}(t) = \frac{1}{\pi(x)} \sum_y \pi(y) p_{yx}(t) = \frac{1}{\pi(x)} (\pi P(t))_x = \frac{1}{\pi(x)} \pi(x) = 1.$$

So if $Z \sim \text{Markov}(\hat{Q})$

$$1 = \sum_y \hat{p}_{xy}(t) = \sum_y \mathbb{P}_x(Z_t = y) = \sum_y \mathbb{P}_x(Z_t = y, t < \zeta) = \mathbb{P}_x(t < \zeta).$$

i.e $\mathbb{P}_x(\zeta > t) = 1$ for all t , so $\mathbb{P}_x(\zeta = \infty) = 1$, i.e non-explosive. □

Definition. Let $X \sim \text{Markov}(Q)$. It is called *reversible* if for all $T > 0$, $(X_t)_{0 \leq t \leq T}$ and $(X_{T-t})_{0 \leq t \leq T}$ have the same distribution.

Definition. A measure λ and a Q -matrix Q are said to be in *detailed balance* if for all x, y

$$\lambda(x)q_{xy} = \lambda(y)q_{yx}.$$

Lemma 2.18. If Q and λ are in detailed balance, then λ is invariant for Q (i.e $\lambda Q = 0$).

Proof.

$$(\lambda Q)_y = \sum_x \lambda(x)q_{xy} = \lambda(y) \sum_x q_{yx} = 0$$

□

Remark. To find an invariant measure, check the detailed balance equation as a first step.

Lemma 2.19. Let $X \sim \text{Markov}(Q)$ be irreducible, non-explosive and π a distribution with $X_0 \sim \pi$. Then π and Q are in detailed balance if and only if $(X_t)_{t \geq 0}$ is reversible.

Proof. X is reversible if and only if $Q = \hat{Q}$ and π is an invariant distribution, where $\hat{q}_{xy} = \frac{\pi(y)}{\pi(x)} q_{yx}$. This happens iff π and Q are in detailed balance. \square

Definition. A birth and death chain X is a continuous time Markov chain on $\mathbb{N} = \{0, 1, \dots\}$ where for $x \geq 1$ $q_{x,x-1} = \mu_x$, $q_{x,x+1} = \lambda_x$, q_{xy} for all other y ; and $q_{01} = \lambda_0$, $q_{0,y} = 0$ for all $y \neq 1$.

Lemma 2.20. A measure π is an invariant measure for a birth and death chain if and only if it solves the detailed balance equation.

Proof. We already have one direction. So we show that if π is invariant it satisfies the detailed balance equation. Indeed, let π be an invariant measure for Q , i.e $\pi Q = 0$. So for all $j \geq 1$,

$$\begin{aligned} (\pi Q)_j &= 0 = \pi_{j-1} q_{j-1,j} + \pi_j q_{j,j} + \pi_{j+1} q_{j+1,j} \\ &= \pi_{j-1} \lambda_{j-1} + \pi_{j+1} \mu_{j+1} - \pi_j (\lambda_j + \mu_j). \end{aligned}$$

So

$$\pi_{j+1} \mu_{j+1} - \pi_j \lambda_j = \pi_j \mu_j - \pi_{j-1} \lambda_{j-1}. \quad (*)$$

For $j = 1$ $(*)$ becomes $\pi_1 \mu_1 - \pi_0 \lambda_0 = 0$. So using induction and plugging in to the RHS of $(*)$ we get

$$\pi_{j+1} \mu_{j+1} = \pi_j \lambda_j.$$

As required. \square

3 Queueing Theory

Queues are processes which can be modelled as customers arriving at a server and then departing.

Q: what is the equilibrium queue length (including customers being served)?

Q: What is the busy period?

Q: Time spent by a customer in the queue/waiting-time (including the service time)?

We use M/G/K notation. The ‘M’ stands for “Markovian arrival” - customers arrive according to a Poisson process of rate λ . The ‘G’ stands for “general distribution” - it is the (iid) service time distribution, if ‘M’ is used instead of ‘G’ this represents $\text{Exp}(\mu)$ service times. The ‘K’ stands for the number of servers ($k = 1$ or ∞).

Let X_t be the queue length at time t (including the customers being served). Then $(X_t)_{t \geq 0}$ is a continuous time process on state space $I = \{0, 1, 2, \dots\}$. If we have a M/M/1 or M/M/ ∞ process, then $(X_t)_{t \geq 0}$ is Markov and in particular it's a birth & death chain with

$$\begin{aligned} \text{M/M/1: } & q_{i,i+1} = \lambda, \quad q_{i,i-1} = \mu \\ \text{M/M}/\infty: & q_{i,i+1} = \lambda, \quad q_{i,i-1} = i\mu \end{aligned}$$

M/M/1:

Theorem 3.1. *Let $\rho = \lambda/\mu$. Then the queue length X (for a M/M/1 process) is transient if and only if $\rho > 1$, recurrent if and only if $\rho \leq 1$ and positive recurrent if and only if $\rho < 1$. In the positive recurrent case, the invariant distribution is*

$$\pi(n) = (1 - \rho)\rho^n, \quad n = 0, 1, \dots$$

And if $\rho < 1$, and $X_0 \sim \pi$, then the wait time (including service time) for a customer that arrives at time t is $\text{Exp}(\mu - \lambda)$.

Proof. The jump chain Y is given by $p_{i,i+1} = \lambda/(\lambda + \mu)$ and $p_{i,i-1} = \mu/(\lambda + \mu)$. This is just a biased SRW on \mathbb{N} (with reflection at 0). Thus Y (and hence X) is transient if $\lambda > \mu$, and recurrent if $\lambda \leq \mu$.

It is non-explosive since $\sup_i q_i = (\lambda + \mu) < \infty$. Thus we have positive recurrence iff there is an invariant distribution. Since X is a birth & death chain, a measure is invariant iff it satisfies detailed balance. Thus $\pi(n)\lambda = \pi(n+1)\mu$, i.e. $\pi(n+1) = \pi(0)(\lambda/\mu)^{n+1}$. So π is normalisable iff $\lambda/\mu = \rho < 1$. When $\rho < 1$, $\pi(n) = (1 - \rho)\rho^n$ is an invariant distribution. So π is the distribution of a (shifted) geometric random variable, i.e. π is the distribution of $Z - 1$ where $Z \sim \text{Geo}(1 - \rho)$.

If $\rho < 1$ and $X_0 \sim \pi$ then $X_t \sim \pi$ (as X is recurrent, π invariant iff $\pi P(t) = \pi$ for all t). So the wait time W of a customer arriving at time t is $W = \sum_{i=1}^{X_t+1} T_i$ where $T_i \sim \text{Exp}(\mu)$ are iid and independent of X_t . As $X_t + 1 \sim \text{Geo}(1 - \rho)$ is independent of $(T_i)_{i \geq 1}$ we have $W \sim \text{Exp}(\mu(1 - \rho)) = \text{Exp}(\mu - \lambda)$ (by Example Sheet 1).

We have expected queue length at equilibrium

$$\mathbb{E}_\pi X_t = \mathbb{E}_\pi Z - 1 = \frac{1}{1 - \rho} - 1 = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}.$$

□

M/M/ ∞ :

Theorem 3.2. *The queue length X_t is positive recurrent for all $\mu > 0$, $\lambda > 0$ with invariant distribution $\text{Poi}(\rho)$ where $\rho = \lambda/\mu$.*

Proof. As X is a birth & death process, we just solve the detailed balance equation:

$$\lambda\pi_{n-1} = n\mu\pi_n \implies \pi_n = \frac{1}{n} \frac{\lambda}{\mu} \pi_{n-1} = \dots = \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n \pi_0.$$

This is always normalisable with $\pi_n = e^{-\lambda/\mu} (\lambda/\mu)^n \frac{1}{n!}$ i.e $\pi \sim \text{Poi}(\rho)$.

We will in fact show Y is positive recurrent. Define $\mu_i = \pi_i q_i$. Then μ is an invariant measure for Y . It is enough to check that μ is normalisable. We have

$$\mu_i = (i\mu + \lambda) e^{-\rho} \frac{\rho^i}{i!} = \rho\mu \left(e^{-\rho} \frac{\rho^{i-1}}{(i-1)!} (i + \rho) \right)$$

and

$$\sum_{i=0}^{\infty} \frac{\rho^{i-1}}{(i-1)!} (i + \rho) = \sum_{i=1}^{\infty} \frac{\rho^{i-1}}{(i-1)!} + \sum_{i=0}^{\infty} \frac{\rho^i}{i!} < \infty$$

so we are done. \square

Let A and D denote the arrival and departure processes associated with a queue (i.e A_t and D_t are the number of customers that have arrived/departed by time t respectively). A, D are increasing processes, and A increases by 1 if and only if X increases by 1; D increases by 1 if and only if X decreases by 1. So $X_t = X_0 + A_t - D_t$. A is a Poisson process of time λ .

Remark. A Poisson process does not have an invariant distribution, but still has the following time-reversing property: if N is a Poisson Process of rate λ , then for any $T > 0$, $\hat{N}_t = N_T - N_{T-t}$ is again a Poisson Process of rate λ on $[0, T]$. Indeed, conditioning on $N_T = n$, the distribution of the jump times is $\frac{n!}{T^n} \mathbb{1}(0 \leq t_1 \leq t_2 \leq \dots t_n \leq T)$.

Theorem 3.3 (Burke's Theorem). *Consider an M/M/1 queue with $\mu > \lambda > 0$ or an M/M/ ∞ queue with $\mu, \lambda > 0$. At equilibrium (i.e $X_0 \sim \pi$), D is a Poisson process of rate λ and X_t is independent of $(D_s : s \leq t)$.*

Remark. This roughly says that “the output of a stationary M/M/ k queue is again a Poisson process”.

Remark. $X_0 \sim \pi$ is essential. Suppose that $X_0 = 5$ for an M/M/1, the first departure happens at $\text{Exp}(\mu)$ and not $\text{Exp}(\lambda)$.

Remark. The processes $(X_s, s \leq t)$ and $(D_s : s \leq t)$ are not independent - clearly D has a jump of +1 exactly when X has a jump of -1.

Proof of Burke's Theorem. As X is a birth & death process, π satisfies the detailed balance equation, i.e if $X_0 \sim \pi$ then X is reversible. Thus for a fixed $T > 0$, with $\hat{X}_t = X_{T-t}$ we have $(\hat{X}_t)_{0 \leq t \leq T} \stackrel{d}{=} (X_t)_{0 \leq t \leq T}$. Hence the arrival process \hat{A} for \hat{X} (until time T) is a Poisson Process of rate λ . But $\hat{A}_t = D_T - D_{T-t}$.

Since the time reversal of a Poisson Process on $[0, T]$ is again a Poisson Process on $[0, T]$, this implies $(D_t)_{0 \leq t \leq T}$ is a Poisson Process of rate λ on $[0, T]$. Since $T > 0$ is arbitrary, this determines the finite-dimensional distributions of D and hence determines the distribution of D , i.e D is a Poisson Process of rate λ on \mathbb{R} .

Independence: as X_0 is independent of $(A_s : 0 \leq s \leq T)$, for the \hat{X} , \hat{X}_0 is independent of (\hat{A}_s) , i.e X_T is independent of $(D_t)_{0 \leq t \leq T}$. \square

Queues in tandem

Suppose that there is an M/M/1 queue with parameters λ and μ_1 . After a customer is served, they immediately join a second M/M/1 queue with parameters λ and μ_2 . Let X and Y denote the queue lengths of the two queues respectively. For (X, Y) have state space $I = \mathbb{N} \times \mathbb{N}$ and the rates are

$$(m, n) \rightarrow \begin{cases} (m+1, n) & \text{with rate } \lambda \\ (m-1, n+1) & \text{with rate } \mu_1 \text{ if } m \geq 1 \\ (m, n-1) & \text{with rate } \mu_2 \text{ if } n \geq 1 \end{cases}$$

Theorem 3.4. (X, Y) is positive recurrent if and only if $\lambda < \mu_1$ and $\lambda < \mu_2$. In this case, the invariant distribution is given by

$$\pi(m, n) = (1 - \rho_1)\rho_1^m(1 - \rho_2)\rho_2^n \text{ where } \rho_1 = \lambda/\mu_1, \rho_2 = \lambda/\mu_2.$$

i.e at equilibrium, X_t and Y_t are independent (for fixed t , not as processes).

Proof 1. Directly check that $\pi Q = 0$. As the rates are bounded, (X, Y) is non-explosive. \square

Proof 2. Note the marginal X is an M/M/1 queue. Thus X is positive recurrent if and only if $\lambda < \mu_1$ with invariant distribution $\pi^1(m) = (1 - \rho_1)\rho_1^m$. By Burke's theorem, if $X_0 \sim \pi^1$, then the departure process of the first queue is a Poisson Process of rate λ , which is the arrival process for the second queue.

So the second queue is M/M/1(λ, μ_2) with invariant distribution $\pi^2(n) = (1 - \rho_2)\rho_2^n$ if $\lambda < \mu_2$. If $X_0 \sim \pi^1$ and $Y_0 \sim \pi^2$ are independent, then $X_t \sim \pi^1$ (as X is recurrent) and also by Burke's theorem, X_t is independent of the departure process until time t , and also independent of Y_0 , so X_t is independent of Y_t .

Also $Y_t \sim \pi^2$ (as Y is recurrent), so $(X_t, Y_t) \sim \pi$. i.e $(X_0, Y_0) \sim \pi \Rightarrow (X_t, Y_t) \sim \pi$ for all t . So π is invariant for (X, Y) (by the following exercise). \square

Exercise: if Z is irreducible, π a distribution and $\pi P(t) = \pi$ for all t , then π is invariant for Z (consider the discrete-time chain $Z_n = (Z_n)$).

Jackson's Network

Have a network of N single-server queues with arrival rates λ_k and service rates μ_k , $1 \leq k \leq N$. After service, each customer in queue i moves to queue j with probability p_{ij} , or exits the system with probability $p_{i0} = 1 - \sum_{j=1}^N p_{ij}$.

We assume $p_{ii} = 0$ and $p_{i0} > 0$ for all $1 \leq i \leq N$. Also assume the system is irreducible, i.e a customer arriving in queue i has a positive probability of visiting queue j at a later time for all $i \neq j$. Thus $I = \{0, 1, 2, \dots\}^N$, where if $x = (x_1, \dots, x_N)$ then x_i is the number of customers in queue i .

If $n = (n_1, \dots, n_N) \in I$ and $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ has all entries 0 except i th entry 1, then

$$\begin{aligned} q_{n, n+e_i} &= \lambda_i \text{ for } i = 1, 2, \dots, N \\ q_{n, n-e_i+e_j} &= \mu_i p_{ij} \text{ for } i, j = 1, \dots, N, n_i \geq 1, i \neq j \\ q_{n, n-e_i} &= \mu_i p_{i0} \text{ for } i = 1, \dots, N, n_i \geq 1 \end{aligned}$$

Definition. We say a vector $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_N)$ satisfies the *traffic equation* if for all $1 \leq i \leq N$

$$\bar{\lambda}_i = \lambda_i + \sum_{\substack{j=1 \\ j \neq i}}^N \bar{\lambda}_j p_{ji}. \quad (*)$$

Remark. $\bar{\lambda}_i$ is the “effective arrival rate” at queue i .

Lemma 3.5. *There exists a unique solution to (*).*

Proof. Uniqueness: see Example sheet 3.

Existence: let $p_{00} = 1$. Then $P = (p_{ij})_{i,j=0}^N$ is a stochastic matrix corresponding to a discrete-time Markov chain (Z_n) . Then (Z_n) is absorbing at 0, so the communicating class $\{1, \dots, N\}$ is not closed, so is transient. Thus if $V_i = \# \text{visits to state } i \text{ by } Z$, then starting from Z_0 , $\mathbb{E}V_i < \infty$ for all $i = 1, \dots, N$.

Let $\mathbb{P}(Z_0 = i) = \frac{\lambda_i}{\lambda}$, for $i = 1, \dots, N$, $\lambda = \sum_{i=1}^N \lambda_i$. Then for all $1 \leq i \leq N$

$$\begin{aligned} \mathbb{E}V_i &= \mathbb{E} \sum_{n=0}^{\infty} \mathbb{1}(Z_n = i) \\ &= \mathbb{P}(Z_0 = i) + \sum_{n=0}^{\infty} \mathbb{P}(Z_{n+1} = i) \\ &= \mathbb{P}(Z_0 = i) + \sum_{n=0}^{\infty} \sum_{j=1}^N \mathbb{P}(Z_n = j) p_{ji} \\ &= \frac{\lambda_i}{\lambda} + \sum_{j=1}^N p_{ji} \sum_{n=0}^{\infty} \mathbb{P}(Z_n = j) \\ &= \frac{\lambda_i}{\lambda} + \sum_{j=1}^N p_{ji} \mathbb{E}V_j \end{aligned}$$

Multiplying throughout by λ and setting $\bar{\lambda}_i = \lambda \mathbb{E}V_i$ we get $\bar{\lambda}_i = \lambda_i + \sum_{j=1}^N \bar{\lambda}_j p_{ji}$. \square

Theorem 3.6 (Jackson, 1957). *Assume that the traffic equation (*) has solution $\bar{\lambda}_i$ such that $\bar{\lambda}_i < \mu_i$ for all $i = 1, \dots, N$. Then the Jackson Network is positive recurrent with invariant distribution*

$$\pi(n) = \prod_{i=1}^N (1 - \bar{\rho}_i) \bar{\rho}_i^{n_i}, \text{ where } \bar{\rho}_i = \frac{\bar{\lambda}_i}{\mu_i}.$$

At equilibrium, the departure processes (to outside) from each queue form independent Poisson processes with rates $\bar{\lambda}_i p_{i0}$.

Remark. At equilibrium, the queue lengths X_t^i are independent for a fixed time t .

Remark. The equilibrium for Jackson Network is not reversible, but there is “partial reversibility”.

Lemma 3.7 (Partial detailed balance). *Let X be a Markov process on I and π be a measure on I . Assume that for each $x \in I$, there is a partition of $I \setminus \{x\}$ as*

$$I \setminus \{x\} = I_1^x \cup I_2^x \cup \dots$$

such that for all $i \geq 1$

$$\sum_{y \in I_i^x} \pi(x) q_{xy} = \sum_{y \in I_i^x} \pi(y) q_{yx}.$$

If π satisfies this, then π is an invariant measure.

Proof. We show $\pi Q = 0$:

$$\begin{aligned} (\pi Q)_y &= \sum_x \pi(x) q_{xy} = \sum_{x \neq y} \pi(x) q_{xy} + \pi(y) q_{yy} \\ &= \sum_i \sum_{x \in I_i^y} \pi(x) q_{xy} + \pi(y) q_{yy} \\ &= \sum_i \sum_{x \in I_i^y} \pi(y) q_{yx} + \pi(y) q_{yy} \\ &= \sum_x \pi(y) q_{yx} \\ &= 0. \end{aligned}$$

□

We are now ready to prove

Theorem 3.8 (Jackson, 1957). *Assume that the traffic equation (*) has solution $\bar{\lambda}_i$ such that $\bar{\lambda}_i < \mu_i$ for all $i = 1, \dots, N$. Then the Jackson Network is positive recurrent with invariant distribution*

$$\pi(n) = \prod_{i=1}^N (1 - \bar{\rho}_i) \bar{\rho}_i^{n_i}, \text{ where } \bar{\rho}_i = \frac{\bar{\lambda}_i}{\mu_i}.$$

At equilibrium, the departure processes (to outside) from each queue form independent Poisson processes with rates $\bar{\lambda}_i p_{i0}$.

Proof. Let $\pi(n) = \prod_{i=1}^N \bar{\rho}_i^{n_i}$. We shall check this satisfies the partial detailed balance equations. Let $A = \{e_i : 1 \leq i \leq N\}$, $D_j = \{e_i - e_j : i \neq j\} \cup \{-e_j\}$ where $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ has all entries 0 except i th entry 1.

When a customer arrives and $n \in I$, $n \rightarrow n + m$ for some $m \in A$. When a customer leaves queue j , $n \rightarrow n + d$ for some $m \in D_j$. Fix n , consider the partition of $I \setminus \{n\}$ given by

$$I \setminus \{n\} = \{n + A\} \cup \bigcup_{j=1}^N \{n + D_j\}.$$

We will show

$$\begin{aligned}\sum_{m \in A} q_{n,n+m} &= \sum_{m \in A} \frac{\pi_{n+m}}{\pi_n} q_{n+m,n}, \\ \sum_{m \in D_j} \pi_n q_{n,n+m} &= \sum_{m \in D_j} \frac{\pi_{n+m}}{\pi_n} q_{n+m,n}.\end{aligned}$$

Note

$$\sum_{m \in D_j} q_{n,n+m} = \mu_j p_{j0} + \sum_{i \neq j} \mu_j p_{ji} = \mu_j$$

and

$$\begin{aligned}\sum_{m \in D_j} \frac{\pi_{n+m}}{\pi_n} q_{n+m,n} &= \frac{\pi_{n-e_j}}{\pi_n} q_{n-e_j,n} + \sum_{i \neq j} \frac{\pi_{n+e_i-e_j}}{\pi_n} q_{n+e_i-e_j,n} \\ &= \frac{1}{\bar{\rho}_j} \lambda_j + \sum_{i \neq j} \frac{\bar{\rho}_i}{\bar{\rho}_j} \mu_i p_{ij} \\ &= \frac{\lambda_j}{\bar{\rho}_j} + \sum_{i \neq j} \frac{\bar{\lambda}_i}{\bar{\rho}_j} p_{ij} \\ &= \frac{\lambda_j + \sum_{i \neq j} \bar{\lambda}_i p_{ij}}{\bar{\rho}_j} \\ &= \frac{\bar{\lambda}_j}{\bar{\rho}_j} \\ &= \mu_j.\end{aligned}$$

Now for A :

$$\sum_{m \in A} q_{n,n+m} = \sum_i \lambda_i$$

and

$$\begin{aligned}\sum_{m \in A} \frac{\pi_{n+m}}{\pi_n} q_{n+m,n} &= \sum_i \frac{\pi_{n+e_i}}{\pi_n} q_{n+e_i,n} = \sum_i \frac{\bar{\lambda}_i}{\mu_i} \mu_i p_{i0} \\ &= \sum_i \bar{\lambda}_i p_{i0} \\ &= \sum_i \bar{\lambda}_i \left(1 - \sum_j p_{ij}\right) \\ &= \sum_i \bar{\lambda}_i - \sum_j \sum_i p_{ij} \bar{\lambda}_i \\ &= \sum_i \bar{\lambda}_i - \sum_j (\bar{\lambda}_j - \lambda_j) \\ &= \sum_i \lambda_i.\end{aligned}$$

Finally as the rates are bounded, it is non-explosive, hence positive recurrent. (Final part of theorem is on the Example Sheet). \square

M/G/1 queue:

Arrival: Poisson process of rate λ . Service time of n th customer: $\xi_n \geq 0$ and (ξ_n) iid with $\mathbb{E}\xi_1 = \frac{1}{\mu}$. Single server.

Denote by $(X_t)_{t \geq 0}$ the queue length, which is no longer a Markov process (service time is no longer memoryless in general).

Let D_n be the departure time of the n th customer. We consider the discrete-time process $Z_n = X(D_n)$.

Proposition 3.9. $Z_n = X(D_n)$, $n = 0, 1, \dots$ is a discrete-time Markov chain with transition matrix

$$\begin{pmatrix} p_0 & p_1 & p_2 & \dots \\ p_0 & p_1 & p_2 & \dots \\ 0 & p_0 & p_1 & p_2 & \dots \\ 0 & 0 & p_0 & p_1 & \dots \\ \vdots & \ddots & \ddots & \ddots & \dots \end{pmatrix}$$

where $p_k = \mathbb{E} \left[e^{-\lambda \xi_1} \frac{(\lambda \xi_1)^k}{k!} \right]$ for $k = 0, 1, \dots$

Proof. Let A_{n+1} be the number of customers arriving after time D_n and during the service time of the $(n+1)$ th customer ξ_{n+1} . Then the A_n are iid (by the independent increment property of a Poisson process), and given ξ_n , $A_n \sim \text{Poi}(\lambda \xi_n)$, i.e. $\mathbb{P}(A_k = k) = \mathbb{E}[\mathbb{P}(A_k = k | \xi_k)] = \mathbb{E} \left[e^{-\lambda \xi_n} \frac{(\lambda \xi_n)^k}{k!} \right] = p_k$.

Now

$$X(D_{n+1}) = \begin{cases} A_{n+1} & \text{if } X(D_n) = 0 \\ X(D_n) + A_{n+1} - 1 & \text{if } X(D_n) > 0 \end{cases}$$

so we have the required transition matrix. \square

Lemma 3.10. Let (Y_i) be iid integer valued random variables and let $S_n = Y_1 + \dots + Y_n$ be the corresponding random walk on \mathbb{Z} starting from 0. If $\mathbb{E}|Y_1| < \infty$, then S is recurrent if and only if $\mathbb{E}Y_1 = 0$.

Proof. Not given. \square

Theorem 3.11. Let $\rho = \frac{\lambda}{\mu}$. If $\rho \leq 1$, the queue is recurrent in the sense that it will hit 0 almost surely. If $\rho > 1$ then it is transient in the sense that there is a positive probability the queue length will never hit 0.

Proof 1. X is transient/recurrent in the sense of the theorem $\iff X(D_n)$ is transient/recurrent in the usual sense. While $X(D_n) > 0$, $(X(D_n))$ is a random walk on \mathbb{Z} with step distribution $Y_i = A_i - 1$. But

$$\mathbb{E}Y_1 = \mathbb{E}A_1 - 1 = \mathbb{E}[\mathbb{E}[A_1|\xi_1]] - 1 = \mathbb{E}[\lambda\xi_1] - 1 = \frac{\lambda}{\mu} - 1 = \rho - 1.$$

If $\rho = 1$ then X is recurrent (by the previous lemma). If $\rho < 1$, then X has a drift to the left, so recurrent. If $\rho > 1$ then X is transient. \square

Proof 2. We will use a hidden branching structure. Say that a customer C_2 is an offspring of C_1 if C_2 arrives during the service of C_1 . This defines a tree. The offspring distribution is iid and distributed as A_1 which given ξ_1 is $\text{Poi}(\lambda\xi_1)$. We have $\mathbb{E}A_1 = \mathbb{E}\mathbb{E}[A_1|\xi_1] = \mathbb{E}[\lambda\xi_1] = \lambda\mathbb{E}\xi_1 = \frac{\lambda}{\mu} = \rho$.

This is a branching process, and we have recurrence (e.g the queue empties out almost surely) if and only if the tree is finite with probability 1, which happens if and only if $\mathbb{E}A_1 = \rho \leq 1$ (see IA Probability). \square

Definition. The time between a customer joining the queue and a customer departing leaving behind an empty queue is called the *busy period*.

Proposition 3.12. For the $M/G/1$ queue with $\lambda < \mu$, the length of the busy period B satisfies

$$\mathbb{E}B = \frac{1}{\mu - \lambda}.$$

Proof. Exercise: use the branching process structure from above. \square

Lemma 3.13. Let $(Y_i)_{i \geq 1}$ be iid \mathbb{Z} -valued random variables and let $S_n = Y_1 + \dots + Y_n$ be the corresponding random walk starting from 0. If $\mathbb{E}|Y_1| < \infty$, then S is recurrent if and only if $\mathbb{E}Y_1 = 0$.

Proof. By the Strong Law of Large Numbers, if $\mathbb{E}Y_1$ exists and is non-zero, $|S_n| \rightarrow \infty$ almost surely.

If $\mathbb{E}Y_1 = 0$ then by the Strong Law of Large Numbers $S_n/n \rightarrow 0$ almost surely. Fix $\varepsilon > 0$. Then for some n large enough

$$\min_{i \leq n} \mathbb{P}(|S_i| \leq \varepsilon n) \geq 1/2. \quad (*)$$

Indeed, choose N_1 large so that for all $n \geq N_1$ have $\mathbb{P}(|S_n| \leq \varepsilon n) \geq 1/2$. Then choose $N_2 > N_1$ large enough so that $\mathbb{P}(|S_i| \leq \varepsilon N_2) \geq 1/2$ for all $i = 1, \dots, N_1 - 1$. Then for $n = N_2$ it holds.

Let

$$\begin{aligned} G_n(x) &= \mathbb{E}_0[\text{\#visits to } x \text{ by time } n] = \mathbb{E}_0 \left[\sum_{k=0}^{\infty} \mathbb{1}(S_k = x) \right] \\ &= \sum_{k=0}^n \mathbb{P}_0(S_k = x). \end{aligned}$$

Clearly, $G_n(x)$ is increasing in n , and for all x , $G_n(x) \leq G_n(0)$ since

$$G_n(x) = \sum_{k=0}^n \mathbb{P}_0(T_x = k) G_{n-k}(0) \leq G_n(0) \sum_{k=0}^n \mathbb{P}_0(T_x = k) \leq G_n(0).$$

Thus taking n as in $(*)$,

$$\begin{aligned}
 (2n\varepsilon + 1)G_n(0) &\geq \sum_{|x| \leq n\varepsilon} G_n(x) = \sum_{|x| \leq n\varepsilon} \sum_{k=0}^n \mathbb{P}(S_k = x) \\
 &= \sum_{k=0}^n \sum_{|x| \leq n\varepsilon} \mathbb{P}(S_k = x) \\
 &= \sum_{k=0}^n \mathbb{P}(|S_k| \leq n\varepsilon) \\
 &\geq \frac{n+1}{2}.
 \end{aligned}$$

So $G_n(0) \geq \frac{1}{4\varepsilon}$, and letting $n \rightarrow \infty$ $\mathbb{E}_0 V_0 \geq \frac{1}{4\varepsilon}$, and since $\varepsilon > 0$ was arbitrary, $\mathbb{E}_0 V_0 = \infty$ so we have recurrence. \square

4 Renewal Processes

Suppose buses arrive every 10 minutes on average, according to a Poisson process of rate $1/10$. How long does one need to wait on average if I arrive at time t ?

What is the “inter-arrival time” that contains t ? It is no longer $\text{Exp}(1/10)$, but larger.

What happens when the n th bus arrives after time ξ_n , where $\xi_n \geq 0$ is iid. Again the length of the interval containing t is larger than ξ_1 . In fact for t large enough, this is the “size-biased” distribution of ξ_1 .

Definition. Let $(\xi_i)_{i \geq 1}$ be iid non-negative random variables, distributed as ξ , with $\mathbb{P}(\xi > 0) > 0$. Set $T_n = \sum_{i=1}^n \xi_i$ and $N_t = \max\{n \geq 0 : T_n \leq t\}$ (the number of renewals until time t for ξ_n the time of the n th renewal). The process $(N_t : t \geq 0)$ is called a *renewal process*.

Remark. If ξ_1, ξ_2, \dots are iid $\text{Exp}(\lambda)$ then (N_t) is a Poisson process of rate λ .

Theorem 4.1. If $\mathbb{E}\xi = \frac{1}{\lambda} < \infty$ then as $t \rightarrow \infty$,

$$\frac{N_t}{t} \rightarrow \lambda \text{ almost surely, and } \frac{\mathbb{E}N_t}{t} \rightarrow \lambda.$$

Remark. We won't prove $\frac{\mathbb{E}N_t}{t} \rightarrow \lambda$ (see Grimett-Strizakel).

Proof. First note that $N_t < \infty$ almost surely and $N_t \rightarrow \infty$ almost surely. Then $T_{N_t} \leq t \leq T_{N_t+1}$. Hence

$$\frac{T_{N_t}}{N_t} \leq \frac{t}{N_t} \leq \frac{T_{N_t+1}}{N_t}.$$

By the Strong Law of Large Numbers, $\frac{T_n}{n} \rightarrow \mathbb{E}\xi = \frac{1}{\lambda}$ and $N_t \rightarrow \infty$ as $t \rightarrow \infty$ almost surely, so $\frac{T_{N_t}}{N_t} \rightarrow \frac{1}{\lambda}$ almost surely and $\frac{T_{N_t+1}}{N_t} = \frac{T_{N_t+1}}{N_t+1} \frac{N_t+1}{N_t} \rightarrow \frac{1}{\lambda}$ almost surely. Thus $\frac{t}{N_t} \rightarrow \frac{1}{\lambda}$ almost surely. \square

Size-biased picking

Now suppose $\mathbb{P}(\xi_1 > 0) = 1$. Let $S_i = \xi_1 + \dots + \xi_i$, $1 \leq i \leq n$. Use S_i/S_n , $1 \leq i \leq n$ to produce a partition of $[0, 1]$ into n subintervals of lengths $Y_i = \xi_i/S_n$. Let U be a uniform $[0, 1]$ random variable independent of ξ_1, \dots, ξ_n , and let \hat{Y} denote the length of the interval containing U .

Since Y_1, \dots, Y_n are identically distributed and $\mathbb{E}[Y_1 + \dots + Y_n] = 1$ so $\mathbb{E}[Y_i] = 1/n$ for all i . What is the distribution of \hat{Y} ? It is not the same as Y_1 since U tends to fall in bigger intervals.

Proposition 4.2. $\mathbb{P}(\hat{Y} \in dy) = ny\mathbb{P}(Y_1 \in dy)$. Formally, $f_{\hat{Y}}(y) \propto yf_{Y_1}(y)$.

Proof.

$$\begin{aligned} \mathbb{P}(\hat{Y} \in dy) &= \sum_{i=1}^n \mathbb{P}\left(\hat{Y} \in dy, \frac{S_{i-1}}{S_n} \leq U \leq \frac{S_i}{S_n}\right) \\ &= \sum_{i=1}^n \mathbb{P}\left(\frac{\xi_i}{S_n} \in dy, \frac{S_{i-1}}{S_n} \leq U \leq \frac{S_i}{S_n}\right) \\ &= \sum_{i=1}^n \mathbb{P}\left(\frac{S_{i-1}}{S_n} \leq U \leq \frac{S_i}{S_n}\right) \mathbb{P}\left(\frac{\xi_i}{S_n} \in dy\right) \\ &= \sum_{i=1}^n y \mathbb{P}\left(\frac{\xi_i}{S_n} \in dy\right) \\ &= ny \mathbb{P}\left(\frac{\xi_1}{S_n} \in dy\right) \\ &= ny \mathbb{P}(Y_1 \in dy). \end{aligned}$$

□

Definition. Let X be a non-negative random variable with distribution μ and $\mathbb{E}X = m < \infty$. Then the *size-biased distribution* of μ is $\hat{\mu}(dy) = \frac{y\mu(dy)}{m}$. We write \hat{X} for a random variable with distribution $\hat{\mu}$.

Remark. $\mathbb{E}\hat{X} = \frac{\mathbb{E}X^2}{\mathbb{E}X} \geq \mathbb{E}X$ if $\mathbb{E}X^2 < \infty$.

Examples.

1. If $X \sim U((0, 1))$ then \hat{X} has distribution $\hat{\mu}(dx) = 2xdx$ on $[0, 1]$.
2. If $X \sim \text{Exp}(\lambda)$ then $\hat{\mu}(dx) = \frac{1}{1/\lambda} \lambda e^{-\lambda x} dx = \lambda^2 x e^{-\lambda x}$ so $\hat{X} \sim \Gamma(2, \lambda)$, i.e. \hat{X} has the same distribution as $X_1 + X_2$ where X_1, X_2 are iid copies of X .

Equilibrium theory of renewal processes

Given a renewal process $(N_t)_{t \geq 0}$ and a time $t > 0$, define

$A(t) = t - T_{N_t}$ the *age process* (time since last renewal)

$E(t) = T_{N_t+1} - t$ the *excess/residual life* (time until next renewal)

$L(t) = T_{N_t+1} - T_{N_t} = A(t) + E(t)$ the length of the current renewal.

What is the distribution of $L(t)$ for t large? Not ξ , but a size-biasing occurs: a big renewal interval is more likely to contain t .

Definition. A random variable is called *arithmetic* if $\mathbb{P}(\xi \in k\mathbb{Z}) = 1$ for some $k > 1$, $k \in \mathbb{Z}$, and *non-arithmetic* if it is not arithmetic, i.e for all $k > 1$, $\mathbb{P}(\xi \in k\mathbb{Z}) < 1$.

Theorem 4.3. Let ξ be non-arithmetic and let $\hat{\xi}$ have the size-biased distribution of ξ . Let $\mathbb{E}\xi = \frac{1}{\lambda}$. Then

$$(L(t), E(t)) \xrightarrow{d} (\hat{\xi}, U\hat{\xi})$$

where $U \sim U((0, 1])$ is independent of $\hat{\xi}$. Similarly

$$(L(t), A(t)) \xrightarrow{d} (\hat{\xi}, U\hat{\xi}).$$

i.e $\mathbb{P}(L(t) \leq x, E(t) \leq y) \rightarrow \mathbb{P}(\hat{\xi} \leq x, U\hat{\xi} \leq y)$ for all x, y .

Remark. $L(t)$ for large t has the size-biased distribution $\hat{\xi}$ and given $L(t)$, the point t falls uniformly within the renewal interval.

Remark. $\mathbb{P}(U\hat{\xi} \leq y) = \lambda \int_0^y \mathbb{P}(\xi > z) dz$. Indeed,

$$\begin{aligned} \mathbb{P}(U\hat{\xi} \leq y) &= \int_0^1 \mathbb{P}(\hat{\xi} \leq y/u) du = \int_0^1 \int_0^{y/u} \mathbb{P}(\hat{\xi} \in dx) du \\ &= \int_0^1 \int_0^{y/u} \lambda x \mathbb{P}(\xi \in dx) du \\ &= \int_0^\infty \lambda x \mathbb{P}(\xi \in dx) \int_0^{y/x \wedge 1} du \\ &= \int_0^\infty \lambda (x \wedge y) \mathbb{P}(\xi \in dx). \end{aligned}$$

Again,

$$\begin{aligned} \lambda \int_0^y \mathbb{P}(\xi > z) dz &= \lambda \int_0^y \int_z^\infty \mathbb{P}(\xi \in dx) dz \\ &= \lambda \int_0^\infty \mathbb{P}(\xi \in dx) \int_0^{x \wedge y} dz \\ &= \lambda \int_0^\infty (x \wedge y) \mathbb{P}(\xi \in dx). \end{aligned}$$

As required.

Example. If $\xi \in \text{Exp}(\lambda)$, then $\hat{\xi} \sim \Gamma(2, \lambda)$ and $\mathbb{P}(U\hat{\xi} \leq y) = \lambda \int_0^y \mathbb{P}(\xi > z) dz = \lambda \int_0^y e^{-\lambda z} dz = 1 - e^{-\lambda y}$ so $U\hat{\xi} \sim \text{Exp}(\lambda)$. Indeed, $E_t \sim \text{Exp}(\lambda)$ for all t .

Example. $\xi \in U((0, 1))$ and $E_t \xrightarrow{d} E_\infty$, then for $0 \leq y \leq 1$,

$$\begin{aligned}\mathbb{P}(E_\infty \leq y) &= \lambda \int_0^y \mathbb{P}(\xi > z) dz \\ &= 2 \int_0^y (1 - z) dz = 2 \left(y - \frac{y^2}{2} \right).\end{aligned}$$

Now we prove

Theorem 4.4. *Let ξ be non-arithmetic and let $\hat{\xi}$ have the size-biased distribution of ξ . Let $\mathbb{E}\xi = \frac{1}{\lambda}$. Then*

$$(L(t), E(t)) \xrightarrow{d} (\hat{\xi}, U\hat{\xi})$$

where $U \sim U((0, 1])$ is independent of $\hat{\xi}$. Similarly

$$(L(t), A(t)) \xrightarrow{d} (\hat{\xi}, U\hat{\xi}).$$

i.e $\mathbb{P}(L(t) \leq x, E(t) \leq y) \rightarrow \mathbb{P}(\hat{\xi} \leq x, U\hat{\xi} \leq y)$ for all x, y .

Proof. We only prove the theorem for ξ discrete, i.e ξ takes values in $\{1, 2, \dots\}$ and time is discrete.

First we prove the convergence for $(E(t))$. $(E(t) : t = 0, 1, \dots)$ is a discrete-time Markov chain on $I = \{1, 2, 3, \dots\}$ with transition probabilities $p_{i,i-1} = 1$ for $i \geq 2$ and $p_{1,n} = \mathbb{P}(\xi = n)$ for all $n \geq 1$.

$(E(t))$ is irreducible, recurrent, and aperiodic (as ξ is non-arithmetic). To get an invariant distribution we solve $\pi = \pi P$, i.e

$$\pi_n = \pi_{n+1} + \pi_1 \mathbb{P}(\xi = n) \quad \forall n \geq 1.$$

hence $\pi_1 = \pi_2 + \pi_1 \mathbb{P}(\xi = 1)$ so $\pi_2 = \pi_1 \mathbb{P}(\xi > 1)$. Then $\pi_2 = \pi_3 + \pi_1 \mathbb{P}(\xi = 2)$, so $\pi_3 = \pi_1 \mathbb{P}(\xi > 2)$. By induction, $\pi_n = \pi_1 \mathbb{P}(\xi > n - 1)$.

Since $\sum_{n \geq 1} \mathbb{P}(\xi > n - 1) = \mathbb{E}\xi = \frac{1}{\lambda}$ we get $\pi_n = \lambda \mathbb{P}(\xi > n - 1)$ for all $n \geq 1$. Thus, $\mathbb{P}(E(t) = y) \rightarrow \pi(y)$ for all $y \geq 1$ (since the chain is irreducible, aperiodic, positive recurrent). Hence for any integer,

$$\mathbb{P}(E(t) \leq y) \rightarrow \sum_{i=1}^y \pi_i = \lambda \sum_{i=1}^y \mathbb{P}(\xi > i - 1) = \lambda \int_0^y \mathbb{P}(\xi > x) dx.$$

Now $((L(t), E(t)) : t \geq 0)$ is also a discrete Markov chain on $I = \{(n, k) : 1 \leq k \leq n\} \subseteq \mathbb{N}^2$ with transition probabilities $p_{(n,k) \rightarrow (n,k-1)} = 1$ for $k \geq 2$, $p_{(n,1) \rightarrow (k,k)} = \mathbb{P}(\xi = k)$ for $k \geq 1$ (which is independent of n). This chain is again irreducible, recurrent, aperiodic, with invariant measure satisfying

$$\pi(n, k - 1) = \pi(n, k) \quad \forall 2 \leq k \leq n \text{ and}$$

$$\pi(k, k) = \underbrace{\sum_{m=1}^{\infty} \pi(m, 1)}_C \mathbb{P}(\xi = k).$$

Then $\pi(n, k) = C \mathbb{P}(\xi = n)$. Since

$$1 = \sum_{n=1}^{\infty} \sum_{k=1}^n \pi(n, k) = \sum_{n=1}^{\infty} C n \mathbb{P}(\xi = n) = C \mathbb{E}(\xi) = \frac{C}{\lambda}$$

have

$$\pi(n, k) = \lambda \mathbb{P}(\xi = n) = \underbrace{\lambda n \mathbb{P}(\xi = n)}_{\mathbb{P}(\hat{\xi}=n)} \underbrace{\frac{1}{n} \mathbb{1}_{\{1 \leq k \leq n\}}}_{\substack{\text{Given } \hat{\xi}=n, E_\infty \text{ is} \\ \text{uniform on } \{1, 2, \dots, n\}}}$$

and $\mathbb{P}(L(t) \leq x, E(t) \leq y) \rightarrow \sum_{n \leq x} \sum_{k \leq y} \pi(n, k)$. \square

Remark. In fact, for any fixed $t > 0$

$$\mathbb{P}(L(t) \geq x) \geq \mathbb{P}(\xi \geq x), \quad \forall x$$

i.e $L(t)$ *stochastically dominates* ξ and hence $\mathbb{E}L(t) \geq \mathbb{E}\xi$. This is called the ‘inspection paradox’.

Renewal-reward process

On top of the renewal structure, suppose there is a reward associated to each renewal, which could be a function of the renewal.

Let (ξ_i, R_i) be iid pairs of random variables (where ξ_i and R_i need not be independent) with $\xi \geq 0$ and $\mathbb{E}\xi = \frac{1}{\lambda} < \infty$. Let $(N_t)_{t \geq 0}$ be the renewal process associated with the ξ_i and let

$$R(t) = \sum_{i=1}^{N_t} R_i$$

be the total reward up to time t . Then

Proposition 4.5. *If $\mathbb{E}|R| < \infty$, as $t \rightarrow \infty$*

$$\frac{R(t)}{t} \rightarrow \lambda \mathbb{E}R \text{ almost surely, and } \frac{\mathbb{E}R(t)}{t} \rightarrow \lambda \mathbb{E}R.$$

Also, for the current reward $\gamma(t) = \mathbb{E}R_{N_t+1}$ we have

Theorem 4.6. *As $t \rightarrow \infty$*

$$\gamma(t) \rightarrow \lambda \mathbb{E}(R\xi).$$

Remark. The factor ξ comes from size-biasing.

Example. Alternating renewal process: a machine runs and then breaks after time X_i when it takes time Y_i to get fixed ($X_i, Y_i \geq 0$ iid). Thus $\xi_i = X_i + Y_i$ is the length of a full cycle and defines a renewal process (N_t) . What is the fraction of time the machine runs in the long-run? If we let $R_i = X_i$ then have

$$\frac{R(t)}{t} \rightarrow \lambda \mathbb{E}X_1 = \frac{\mathbb{E}X_1}{\mathbb{E}X_1 + \mathbb{E}Y_1} \text{ almost surely.}$$

What is the probability $p(t)$ that the machine is on at time t ? $\mathbb{E}R(t) = \int_0^t P(s)ds$ and we expect (and is true under suitable assumptions that)

$$p(t) \rightarrow \frac{\mathbb{E}X_1}{\mathbb{E}X_1 + \mathbb{E}Y_1}.$$

Remark. $\frac{R(t)}{t}$ is not exactly the same as before, as the reward in reward-renewal processes are only collected at the end of the cycle. But in the long-run this won't change the answer for the limit of $R(t)/t$.

Example. Busy period for M/G/1: assume $\rho < 1$. Let I_n and B_n denote the lengths of the n th idle and busy periods respectively. Then (B_n, I_n) is an alternating reward-renewal process. Hence if $p(t)$ is the probability the server is idle at time t , we have

$$p(t) \xrightarrow{t \rightarrow \infty} \frac{\mathbb{E}I_1}{\mathbb{E}I_1 + \mathbb{E}B_1} = \frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + \frac{1}{\mu - \lambda}} = \frac{\mu - \lambda}{\mu} = 1 - \frac{\lambda}{\mu}.$$

Since $I_n \sim \text{Exp}(\lambda)$ by the Markov property and $\mathbb{E}B_1 = \frac{1}{\mu - \lambda}$ by an earlier result. For an M/M/1 queue we would have

$$p(t) \xrightarrow{t \rightarrow \infty} \pi_0 = 1 - \rho$$

.