1 Markov Chains

Definition of Markov Chain $P(X_{n+1}=j|\{X_i=k_i\}) = P(X_{n+1}=j|X_n=i) = p_{ij}$
$P(X_1 = i_1, \ldots, X_n = i_n|X_0 = i_0) = \prod_{k=0}^{n-1} p_{k,k+1}$
$\exists\{Z_n\}$iid taking values in E and $h : S \times E \to S$, $X_n = h(X_{n-1}, Z_n)$ $p_{ij} = P(h(i, Z_1) = j)$

Higher order transition probabilities
Distribution μ_0 on S, distribution of $X_n \sim \mu_n = \mu_0 P^n$. Entries in P^n are $P_i(X_n = j) = p_{ij}^{(n)}$
$P_\mu(X_n = j) = \sum \mu_i p_{ij}^{(n)}$
$P^{n+m} = P^n P^m \Rightarrow p_{ij}^{n+m} = \sum p_{ik}^{(n)} p_{kj}^{(m)} \Rightarrow p_{ij}^{(n+m)} \geq p_{ik}^{(n)} p_{kj}^{(m)} \forall k$

Hitting times and return times
$\tau_A = \inf\{n \geq 0 \,|\, X_n \in A\}$ $\eta_j = \inf\{n \geq 1 \,|\, X_n = j\}$ If $i \neq j$ $P_i(\tau_j \in B) = P_i(\eta_j \in B)$
Successive returns to j $\tau_j(0) = 0$, $\tau_j(1) = \eta_j$, $\tau_j(n) = \inf\{k > \tau_j(n-1) \,|\, X_k = i\}$ given $\tau_j(n-1) < \infty$
Dissection principle principle says $P_i(\cap\{\tau_j(k) - \tau_j(k-1) = m_k, (X_{\tau_j(k-1)+1}, \ldots, X_{\tau_i(k)}) \in A_k\}) = P_i(\tau_j(1) = m_1, (X_1, \ldots, X_n) \in A_i) \prod P_j(\tau_j(1) = m_k, (X_1, \ldots, X_{m_k}) \in A_k)$
Implications:
WRT $P_j(\cdot|\{\tau_j(i) < \infty\}_i)$ the collection $\{\tau_j(k) - \tau_j(k-1) \,|\, k = 1:n\}$ are iid
$\forall n, m \in \mathbb{N}, i.j.k.\{r_l\} \in S$ $P_i(X_{\tau_j(n)+m} = k|X_l = r_l,\ 1 \leq l \leq \tau_j(n), X_{\tau_j(n)} = j) = P_i(X_{\tau_j(n)+m} = k|X_{\tau_j(n)} = j) = p_{jk}^{(m)}$
$\forall i, j, \{k_n\} \in S, n, m \in \mathbb{N}, P_i(\tau_j(1) = m, \{X_{\tau_j(1)+l} = k_l\}_l) = P_i(\tau_j(1) = m) P_j(\{X_l = k_l\}) = P_i(\tau_j(1) = m) p_{jk_1} \prod p_{k_l k_{l+1}}$

Communicating classes of states
j accessible from i $(i \to j)$ if $P_i(\tau_j < \infty) > 0$, & $\exists n \in \mathbb{Z}^+ \ni p_{ij}^{(n)} > 0$. $C_i = \{j \,|\, i \leftrightarrow j\}$
C is closed if: $P_i(\tau_{C^c} = \infty) = 1$, $P_i(X_n \in C) = 1$, $p_{ij} = 0$ for $i \in C$, $j \notin C$
When $p_{jj} = 1$, $C_j = \{j\}$ is closed and absorbing
Chain irreducible when (equivalently): S has a single communicating class, $i \leftrightarrow j$ $\forall i, j \in S$, $\forall\, i, j \in S\, \exists\, n \in \mathbb{Z}^+ \ni p_{ij}^{(n)} > 0$

Transience and recurrence
Notation: $f_{ij}^{(n)} = P_i(\eta_j = n)$, $f_{ij} = P_i(\eta_j < \infty)$, $N_j = \sum_{n\in\mathbb{N}} 1\{X_n = j\}$
$i \in S$ is recurrent when (equivalently): $f_{ii} = 1$, $E_i[N_i] = \sum_{n\in\mathbb{N}} p_{ii}^{(n)} = \infty$, $P_i(N_i = \infty) = 1$
State is positive recurrent if $E_i[\eta_i] < \infty$ and null recurrent when $E_i[\eta_i] = \infty$. If $|S| < \infty$ recurrence $\Rightarrow$ positive recurrence.
$i \in S$ is transient when (equivalently): $f_{ii} < 1$, $E_i[N_i] = \sum_{n\in\mathbb{N}} p_{ii}^{(n)} < \infty$, $P_i(N_i < \infty) = 1$, $E_j[N_i] = \sum_{n\in\mathbb{N}} p_{ji}^{(n)} < \infty\, \forall\, j \in S$.
Properties: If i is recurrent and $i \to j$ then j is recurrent, $i \leftrightarrow j$, $f_{ij} = f_{ji} = 1$, and $C_i = C_j$ is closed. Moreover $k \notin C_i$, $f_{ik} = f_{ki} = 0$, i transient and $i \leftrightarrow j$ then j transient, all states in an irreducible, finite state space MC are positive recurrent
N_j has a geometric distribution. $\forall\, i, j \in S$, $P_i(N_j = 0) = 1 - f_{ij}$, $P_i(N_j = k) = f_{ij} f_{jj}^{k-1}(1 - f_{jj})$ $k \in \mathbb{N}, P_i(N_j = \infty) = f_{ij}$ if recurrent else 0. If j is transient $E_i[N_j] = \frac{f_{ij}}{1-f_{jj}}$

Periodicity
Period of state i $d(i) = \gcd\{n \geq 1 \,|\, p_{ii}^{(n)} > 0\}$ (If $\{n \geq 1 \,|\, p_{ii}^{(n)} > 0\} = \emptyset$, then $d(i) = 1$) If $d(i) = 1$ i is aperiodic, else periodic.
Properties: If $p_{ii} > 0$, $d(i) = 1$, if $i \leftrightarrow j$, then $d(i) = d(j)$

Canonical Decomposition
S can be decomposed as disjoint union $S = T \cup (\cup C_\alpha)$. C_α closed. $\forall\, \alpha$, $p_{ij}^{(n)} = 0\, \forall n \in \mathbb{N}$, $i \in C_\alpha$, $j \notin C_\alpha$ The transition matrix can be written in block form $P = \begin{array}{c} \\ T\{ \\ T^c\{ \end{array} \!\!\begin{array}{c} \overbrace{\quad}^{T} \quad \overbrace{\quad}^{T^c} \\ \begin{pmatrix} Q & R \\ 0 & \tilde{P} \end{pmatrix} \end{array}$ Q is the "fundamental matrix" with properties: $\forall i, j \in T$ $n \in \mathbb{N}$ $P_i(X_n = j) = [Q^n]_{ij}, \sum_{n\in\mathbb{Z}^+} Q^n = (I - Q)^{-1}$.
τ hitting time of T^c, then $[E_i N_j]_{i,j\in T} = (I - Q)^{-1} - I$, $[P_i(X_\tau = j)]_{i\in T, j\in T^c} = (I - Q)^{-1} R$, $[E_i \tau]_{i\in T} = (I - Q)^{-1} \sum_{n\in\mathbb{N}} e_n$

Invariant measures and stationary distributions
An invariant measure is $\{\nu_j\} \geq 0$, $v_j = \sum_{k\in S} \nu_k p_{kj}$ $\forall j \in S$. An invariant distribution $\{\pi_j\}$ is an invariant measure and $\sum_{j\in S} \pi_j = 1$.
Properties: $\{X_n\}$ is stationary WRT P_π $P_\pi((X_n, \ldots, X_{n+k}) \in A) = P_\pi((X_0, \ldots, X_k) \in A) \Rightarrow P_\pi(X_n = j) = \pi_j$, $i \in S$ recurrent, $\nu_j = E_i[\sum_{n=0}^{\eta_i - 1} 1\{X_n = j\}]$ is an invariant measure, i is positive recurrent then $\sum_{j\in S} E_i[\sum_{n=0}^{\eta_i-1} 1\{X_n = j\}] = E_i \eta_i < \infty \Rightarrow \pi_j = \frac{E_i[\sum_{n=0}^{\eta_i-1} 1\{X_n=j\}]}{E_i \eta_i}$ is a stationary distribution, $\{X_n\}$ irreducible, recurrent, and μ, ν are invariant measures then $\exists c > 0 \ni \mu = c\nu$, $\{X_n\}$ is irreducible positive recurrent then $\exists!$ stationary distribution given by $\pi_j = \frac{1}{E_j[\eta_j]}$, $\{X_n\}$ irreducible then TFAE $\exists!$ stationary distribution $\{X_n\}$ is positive recurrent.

Limits in Markov chains
Independence from dissection principle implies SLLN for Markov chains: $\{X_n\}$ irreducible and positive recurrent, $f : S \to [0, \infty)$, and μ is a distribution on S, then $\lim \frac{1}{N} \sum_{n=0}^{N} E_\mu[f(X_n)] = \sum_{j\in S} f(j)\pi_j$.
Let X_n converge in distribution ($p_{ij}^{(n)} \to \pi_j \forall i \in S$), then π is stationary. If $\{X_n\}$ is irreducible, three cases, transient $\Rightarrow$ $\lim p_{ij}^{(n)} = 0\, \forall i, j \in S$, null recurrent and aperiodic $\Rightarrow$ $\lim p_{ij}^{(n)} = 0\, \forall i, j \in S$, ergodic (irreducible, positive recurrent, and aperiodic) $\Rightarrow \lim p_{ij}^{(n)} = \pi_j\, \forall i, j \in S$.

2 Poisson Processes

Definitions
A counting/point process on $[0, \infty)$ is a continuous-time stochastic process. $\{N(t)\}_{t\geq 0}$ with (monotonically increasing $T_0 = 0$) points $\{T_n\}_{n\geq 0}$ where $N(t) = \sum_{n\in\mathbb{N}} 1\{T_n \leq t\} = \#$ of points in $[0, t]$. The increments $N(s, t] = N(t) - N(s) = \sum_{n\in\mathbb{N}} 1\{s < T_n \leq t\} = \#$ of points in $(s, t]$.
$N(t) = \max\{n \geq 0 \,|\, T_n \leq t\}$, $T_n = \min\{t \geq 0 \,|\, N(t) = n\}$, $\{T_n \leq t\} = \{N(t) \geq n\}$
A counting process has independent increments if $\{N(s_i, t_i]\}$ are independent rvs when $\cap(s_i, t_i] = \emptyset$. It has stationary increments if $N(s, t] \stackrel{d}{=} N(0, t - s]\, \forall s < t$.
A counting process $\{N(t)\}$ is a (homogeneous) PP with rate $\lambda > 0$ if $N(0) = 0$, $\{N(t)\}$ has independent increments, and $N(t) \sim Poisson(\lambda t)$ $\forall t \geq 0$. The last part can be replaced with $P(N(t, t+h] = 1) = \lambda h + o(h)$, $P(N(t, t+h] \geq 2) = o(h)$, $h \downarrow 0$, $\forall t \geq 0$.

Properties
$\{N(t)\}$ λ rate HPP with points $\{T_n\}$
$N(s, t] \sim Poisson(\lambda(t - s))$ $\forall s < t$, $\forall n \in \mathbb{N}$ $T_n \sim \Gamma(n, \lambda)$, the inter-arrival times $\{T_{i+1} - T_i\} \sim Exp(\lambda)$ iid rvs, $s < t (N(s)|N(t) = n) \sim Bin(n, \frac{s}{t})$, $\{T_n\}$ satisfies the order statistic property $(\{T_i\}_1^n | N(t) = n) \stackrel{d}{=} (\{T_i\}|T_{n+1} = t) \stackrel{d}{=} (U_{(i)})$ where $\{U_i\}$ are iid $Unif[0, t]$ rvs. In practice this means when f is a symmetric function so that $f(\{U_i\}) = f(\{U_i\})$ then $E[f(\{T_i\}_1^{N(t)})]$ can be computed by conditioning on the value of $N(t)$. For example, $E[\sum_{n=1}^{N(t)} g(T_n)] = E[N(t)]E[g(U_1)] = \lambda \int_0^t g(u)du$ and $E[\prod_{n=1}^{N(t)} h(T_n)] = E[(E[h(U_1)])^{N(t)}] = e^{\lambda(\int_0^t h(u)du - t)}$

Thinning
iid Thinning: Sps $\{L_n\}_{n\in\mathbb{N}}$ are iid natural valued rvs with pmf $\{p_j\}$ independent of $\{N(t)\}$ and $\{T_n\}$ then the thinned process $N_j(t) = \sum_{n\in\mathbb{N}} 1\{T_n \leq t, L_n = j\} = \sum_1^{N(t)} 1\{L_n = j\}$ are independent PP with rates $\lambda\{p_j\}$ respectively.
Time-dependent thinning: Sps $\forall t \geq 0$, $\{p_j(t)\}$ is a pmf and that the distribution of labels $\{L_n\}$ is determined by $P(\{L_i = j_i\}|\{T_l\}) = \prod p_{jk}(T_k)$. Then for each $t \geq 0$ $N_j(t)$ are independent with $N_j(t) \sim Poisson(\lambda \int_0^t p_j(s)ds)$.(In fact, $\{\{N_j(t)\}\}$ are independent nonhomogeneous PP with local intensities $\lambda\{p_j(t)\}$)

Nonhomogeneous Poisson Processes
Definition: A counting process $\{N(t)\}$ is a nonhomogeneous Poisson process with (local) intensity $\alpha : [0, \infty) \to [0, \infty)$ if $N(0) = 0$ wp1, $\{N(t)\}$ has independent increments, $\forall t \geq 0$, $N(t) \sim Poisson(\int_0^t \alpha(s)ds)$. Equivalently, the last one can be replaced by $P(N(t, t+h] = 1) = \alpha(t)h + o(h)$, $P(N(t, t+h] \geq 2) = o(h)$, $h \downarrow 0$, $\forall t \geq 0$.
Time Changes: HPP and NHPP are related by a transformation on $[0, \infty)$. Let $\{N(t)\}$ be a NHPP with intensity $\alpha(t)$ and $\{M(t)\}$ be a HPP with rate 1. Let $\mu(t) = \int_0^t \alpha(s)ds$ and $\mu^{-1}(t) = \inf\{s \geq 0 \,|\, \mu(s) \geq t\}$ Then $\{M(\mu(t))\} \stackrel{d}{=} \{N(t)\}$ and $\{N(\mu^{-1}(t))\} \stackrel{d}{=} \{M(t)\}$

Related Processes
Compound Poisson with parameter λ and component F is the law of $W = \sum_i^N X_i$ where $N \sim Poisson(\lambda)$ independent of $\{X_i\} \sim iid\, F$. Compound Poisson distributions arise in the Poisson processes context by the order statistic property: If $\{N(t)\}$ is a λ rate HPP with points $\{T_n\}$ $W = \sum_1^{N(t)} f(T_n)$ has a compound Poisson distribution with parameter λt and component distribution $f(Unif[0, t])$. If $f(x) = 1_A(x)$ for some $A \subseteq [0, t]$ then $W \sim Poisson(\lambda|A|)$.
A compound PP $\{X(t)\}$ can be represented as $X(t) = \sum_1^{N(t)} \zeta_i$ where $\{N(t)\}$ is a PP independent of the iid rvs $\{\zeta_i\}$
Doubly-stochastic PP: If $\{N(t)\}$ is a PP and $\{\Lambda(t)\}_{t\geq 0}$ is another independent continuous-time non-decreasing process then $\{N(\Lambda(t))\}$ is called a doubly-stochastic PP. A simple type is a conditional PP $\{N(t)\}$ "driven" by a random variable Λ
$(\{N(t)\}|\Lambda = \lambda) \sim \lambda$ rate HPP

Distribution	**PMF/PDF**	**CDF**	**Mean**	**Variance**	**MGF**
Poisson (λ)	$\frac{e^{-\lambda}\lambda^k}{k!}$	$\sum_{i=0}^{k} \frac{e^{-\lambda}\lambda^i}{i!}$	λ	λ	$e^{\lambda(e^t-1)}$
Exponential (λ)	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$ $(t < \lambda)$
Gamma (k, λ)	$\frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}$	$\int_0^x \frac{\lambda^k t^{k-1} e^{-\lambda t}}{\Gamma(k)} dt$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)^k$ $(t < \lambda)$
Geometric (p)	$(1-p)^{k-1}p$	$1-(1-p)^k$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$ $(t < -\ln(1-p))$
Uniform (a, b)	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Bernoulli (p)	$p^k(1-p)^{1-k}$	$\begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$	p	$p(1-p)$	$1-p+pe^t$
Binomial (n, p)	$\binom{n}{k}p^k(1-p)^{n-k}$	$\sum_{i=0}^{k} \binom{n}{i}p^i(1-p)^{n-i}$	np	$np(1-p)$	$(1-p+pe^t)^n$
Normal (μ, σ^2)	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$	μ	σ^2	$e^{\mu t+\frac{1}{2}\sigma^2 t^2}$