MATH 303 Review Sheet:

General Probability rules	Total Law of Probability: $$p(x) = \sum_y p(x,y) = \sum_y p(x\| y) * p(y)$$ Conditional Expectation: $$E(X) = \sum_y E(X\|Y) * p(Y)$$ Notes on some useful distributions: 1. If $X \sim Binom(n,p)$, then $E(X) = np$, $Var(X) = np(1-p)$, $f(x) = \binom{n}{x}(p)^x(1-p)^{n-x}$ 2. If $X \sim Geom1(\pi)$, then $f(x) = (1-\pi)^x\pi$, $E(x) = \frac{1-\pi}{\pi}$, $Var(x) = \frac{1-\pi}{\pi^2}$ 3. If $X \sim Geom0(\pi)$, then $f(x) = (1-\pi)^{x-1}\pi$, $E(x) = \frac{1}{\pi}$, $Var(x) = \frac{1-\pi}{\pi^2}$ Geometric Series: $$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$$ Binomial Theorem: $$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$
Markov Property	For all $n \in$ N & $X_i \in S$, $X_n, n \geq 0$ is a Markov Chain when it satisfies: $$p(X_{n+1} = x_{n+1} \| X_0 = x_0, \dots, X_n = x_n) = p(X_{n+1} = x_{n+1} \| X_n = x_n)$$ Literally, the next state only depends on the current state.
Homogeneous Markov Chain	For all $x, y \in S$, a homogeneous Markov Chain should have $p(X_{n+1} = y \| X_n = x)$ being the same for all states. For Homogenous MC: $E(X_{n+s} \| X_s = i) = E(X_n \| X_0 = i)$, $E(X_n \| X_0 = j, X_i = i) = E(X_{n-1} \| X_0 = i)$
Chapman-Kolmogorov equation	$$p_{i,j}^{m+n} = \sum_k p_{i,k}^m * p_{k,j}^n$$ n-step transition matrix: $\tilde{p}^{(n)} = \tilde{p}^n$
Accessibility & Communication	1. Accessible: S_j is accessible from S_i if there exists $n \epsilon N$ such that $p_{ij}^n > 0$. 2. Communicate: S_j communicates with S_i if S_i is accessible from S_j & S_j is accessible from S_i. - Communication is an equivalent relationship (reflexive: S_i communicate with itself, transitive, symmetric) - All communicating classes partition the state space S. - For an irreducible MC, there is only 1 recurrent communicating class.
Period of States	Period of States $d(i)$: greatest common divisor of $n \in N, p_{i,i}^n > 0$. - All States of the same communicating class share the same period (class property) Aperiodic: a state or MC has a period of 1.
Recurrence & Transience	Let $f_i = p(X_n = i,\ n \geq 1, n \in N \| X_0 = i)$ (probability of return to state i if started from state i): 1. If state i is recurrent, then we have: $f_i = 1$ 2. If state i is transient, we have $f_i < 1$ Proposition: Transience & Recurrence are class properties. Proposition: for an irreducible, finite MC, the MC is recurrent
Number of Returns (Visits) Matrix Inverse:	Let $N_i = \#\ (n \geq 0: X_n = i) \cup \infty$, N_i denotes to number of visits to state i. If i is recurrent, we have:

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. $\tilde{p}^T$: transition matrix for all transient states.	$\begin{cases} p(N_i = \infty \mid X_0 = i) = 1 \\ E(N_i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{i,i}^n = 1 \\ same\ for\ all\ other\ states\ in\ the\ same\ C_i\ with\ i \end{cases}$ If i is transient, we have: $\begin{cases} E(N_i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{i,i}^n = \frac{1}{1-f_i} < \infty \\ f_i < 1 \\ same\ for\ other\ states\ in\ the\ same\ C_i \end{cases}$ $S_{i,j}$: mean Time (Number of Visits) to Transient state j starting from state i $S_{ij} = E(N_j \mid X_0 = i) = \left[\left(I_n - \widetilde{P_T}\right)^{-1}\right]_{ij}$
Closedness	A communicating Class is closed if all $i \in C$ & $j \notin C, p_{ij} = 0$. Proposition: Finite, closed communication class is recurrent; communication class not closed is always transient.
Stirling Approximation	Stirling Approximation: $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ $p_{0,0}^{2n} = \binom{2n}{n} p^n(1-p)^n = \frac{2n!}{n!n!} p^n(1-p)^n \approx \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)^2} p^n(1-p)^n \approx \frac{2\sqrt{\pi n} * 2^{2n} p^n (1-p)^n}{2\pi n} = \frac{2^{2n} p^n (1-p)^n}{\sqrt{\pi n}}$ Proposition: If $\sum p_{0,0}^{2n} < \infty$ (converges), then $\sum \frac{2^{2n} p^n (1-p)^n}{\sqrt{\pi n}} < \infty$ (converges) Random Walk is recurrent when $p = \frac{1}{2}$ & $dim \leq 2$, else it is transient.
Positive Recurrence & Null Recurrence	Given a recurrent state i, let T_i be the first time to revisit i after we started from i, we have: $m_i = E(T_i \mid X_0 = i)$ m_i is the mean time to return to state i. 1. i is positive recurrent if $m_i < \infty$ (class properties, always true when MC is finite & irreducible) 2. i is null recurrent if $m_i = \infty$ (class properties, can only happen with infinite MC)
Stationary Distribution	A vector is called the stationary distribution if it satisfies: (there can be more than 1 stationary distributions) 1. $\pi = \pi\tilde{P}$ 2. $\sum_i \pi_i = 1$ 3. $0 \leq \pi_i \leq 1$ For an irreducible MC, 1. If $\pi = \pi\tilde{P}$ has no solution, then MC is null recurrent or transient. 2. If $\pi = \pi\tilde{P}$ has solution, then MC is positive recurrent. Ergodic: a state or MC is positive recurrent & aperiodic. - For finite state MC, ergodic means recurrent & aperiodic Big Theorem: For an irreducible, ergodic MC (if MC finite then irreducible & aperiodic): 1. There is a unique stationary distribution, i.e., one vector satisfying: $\pi = \pi\tilde{P}$, $\sum_i \pi_i = 1$ (from irreducibility) 2. Limiting distribution is the stationary distribution: $\lim_{n\to\infty} \alpha\tilde{P}^n = \pi$ (from aperiodicity) 3. Mean time needed to return to state i: $m_i = \frac{1}{\pi_i}$ (from irreducibility) 4. $\pi_i = \lim_{n\to\infty} \frac{\#\ of\ visits\ to\ state\ i\ till\ n}{n} = long\ run\ proportion\ of\ time\ spent\ at\ i$ (from irreducibility)
Doubly Stochastic Markov Chain	$\tilde{P}$ is called doubly stochastic if its columns also sum up to 1. If a MC is doubly stochastic, then it has the following stationary distribution: $\pi = (\pi_i \quad \cdots \quad \pi_n) = \left(\frac{1}{n} \quad \cdots \quad \frac{1}{n}\right)$

Time-Reversable Markov Chain	A Markov chain is time reversable if: $$\tilde{Q} = \tilde{P}$$ $$q_{ij} = p_{ji} * \frac{\pi_j}{\pi_i} = p_{ij}$$ Detailed Balance Equation (to prove time reversibility & find stationary distribution): $$x_i p_{ij} = x_j p_{ji}, x = \{x_1 \ldots \ldots x_n\}$$ $$\sum_i x_i = 1$$ If satisfied, x is the stationary distribution & MC is time reversable.
Generating Functions	For a random variable $\xi \in \{1, 2, 3, \ldots\}$, its generation function $G_\xi(s)$ satisfies: $$G_\xi(s) = E(s^\xi) = \sum_{j \geq 0} s^j p(\xi = j)$$ Properties of Generating Functions: 1. If X & Y are independent random variables, then $G_{X+Y}(s) = G_X(s) * G_Y(S)$ 2. Let $T = X_1 + X_2 + \cdots X_N = \sum_{i=1}^{N} X_i$, such that all X_i are independent & follow the same distribution, then suppose N is independent from all X, we have: $$G_T(s) = G_N(G_X(s)) = G_N(s) \circ G_X(s)$$ 3. Implication: $G_{Z_n}(s) = G_n(s) = G_n\big(G_X(s)\big)$ Some Derivations from Generating Functions: 1. $G_X(0) = G(0) = \sum_{j \geq 0} s^j p(X = j), s = 0 = p(X = 0)$ 2. $G_X(1) = G(1) = \sum_{j \geq 0} s^j p(X = j), s = 1 = \sum_{j \geq 0} p(X = j)$ 3. $G'_X(s) = \sum_j j * s^{j-1} p(X = j)$ 4. $G'_X(1) = \sum_j j * s^{j-1} p(X = j), s = 1 = \sum_{j \geq 0} j * 1 * p(X = j) = E(X)$ 5. $G''_X(s) = \sum_j j * (j-1) * s^{j-2} p(X = j)$ 6. $G''_X(1) = \sum_j j * (j-1) * s^{j-2} p(X = j), s = 1 = \sum_j j * (j-1) * 1 * p(X = j) = \sum_j j^2 p(X = j) - \sum_{j \geq 0} j p(X = j) = E(X^2) - E(X)$ Expected Values & Variance of X in relation to its generating function: 1. $E(X) = G_X{}'(1)$ 2. $Var(X) = E(X^2) - \big(E(X)\big)^2 = E(X^2) - E(X) + E(X) - \big(E(X)\big)^2 = G''_X(1) + G'_X(1) - \big(G'_X(1)\big)^2$ How to Find $p(X = k)$: 1. Look at the coefficient before S^k 2. Generally, $p(X = k) = \frac{G_X^{(k)}(0)}{k!}$
Branching Process	Branching Process models $(Z_n)_{n \geq 0}$: the population at generation n with assumption that $Z_0 = 1$. X is a random variable to having the distribution of the number of offspring an individual can have. Properties at the Branching Process: 1. MC is $(Z_n)_{n \geq 0}$ 2. $S = \{0, 1, 2, 3, \ldots.\} \in N$ 3. Absorbing State (recurrent state): $S = 0$ (extinction) 4. Transient states: $S \notin 0$ Generating Function of the Branching Process: 1. $G_{Z_n}(s) = G_n(s) = G\Big(G\Big(G\big(..G(s)\big)\Big)\Big), n\ compositions$ 2. $G_{m+n}(s) = G_m\big(G_n(s)\big) = G_n\big(G_m(s)\big) = G_m(s) \circ G_n(s)$ 3. $p(Z_n = k) = \frac{G_n^{(k)}(0)}{k!} = coefficient\ of\ s^k$ Mean & Variance of Z_n: Let $\mu = E(X), \sigma^2 = Var(x)$, then:

	1. $E(Z_n) = \mu^n$ 2. $Var(Z_n) = \begin{cases} n\sigma^2, & when\ \mu = 1 \\ \frac{\sigma^2(\mu^n-1)\mu^{n-1}}{\mu-1}, & when\ \mu \neq 1 \end{cases}$ Probability of Extinction: 1. Extinction at generation n: $\{Z_n = 0\}$, Extinction by Generation n: $\{Z_n = 0\}$ 3. Probability of Extinction: $p(Z_n = 0)$; Probability of Eventual Extinction: $\lim_{n\to\infty} p(Z_n = 0) = \lim_{n\to\infty} G_n'(0)$ 4. Theorems: $p(eventual\ extinction) = \eta$ should satisfy the following conditions: i. η is the smallest non-negative root of $G(s) = s$ ii. If $\mu < 1$, then $\eta = 1$ (this means on average every individual cannot produce an offspring, thus the total population is shrinking at every generation, which means it will eventual go extinct) iii. If $\mu > 1$, then $\eta < 1$ iv. If $\mu = 1$, then: a. If $\sigma^2 = 0$, then $\eta = 0$. ($Z_n = 1$ for all n) b. If $\sigma^2 > 0$, then $\eta = 1$
Exponential Distribution	We call X an exponential random variable with parameter $\lambda > 0$, $X \sim Expon(\lambda)$ if X has the following density: $f(x) = \begin{cases} \lambda e^{-\lambda x}, x \geq 0 \\ 0, \quad x < 0 \end{cases}$ 1. Distributional Rules of Exponential Random Variables: Suppose $X \sim Expon(\lambda)$ i. $E(X) = \frac{1}{\lambda}$ ii. $Var(X) = \frac{1}{\lambda^2}$ iii. $p(X > t) = \int_t^\infty \lambda e^{-\lambda x} = \lambda * \left(-\frac{1}{\lambda}\right)\left(e^{-\lambda*\infty} - e^{-\lambda*t}\right) = \lambda * \left(-\frac{1}{\lambda}\right) * \left(-e^{-\lambda*t}\right) = e^{-\lambda t}$ 2. Properties of Exponential Distributions: i. Memoryless Property: suppose $T \sim Expon(\lambda)$ a. $p(T \geq t + s \| T \geq s) = p(T \geq t)$ b. $p(T \leq t + s \| T \geq s) = 1 - p(T \geq t + s \| T \geq s) = 1 - p(T \geq t) = p(T \leq t)$ c. Due to the memoryless property, suppose we have 3 people (A, B, C) all being served with time $T \sim Expon(\lambda)$ with two people (A, B) being served first, then suppose A left first, the probability of C leaving before B is $p(T_C < T_B) = \frac{\lambda}{2\lambda} = \frac{1}{2}$ (note that this is because when A left the remaining time of B being served is still $Expon(\lambda)$). ii. Minimum of Exponential: Suppose $X \sim Expon(\lambda_1), Y \sim Expon(\lambda_2)$, such that X & Y are independent, $Z = \min(X, Y)$, then we have: $p(Z \geq t) = p(\min(X, Y) \geq t) = p(X > t, Y > t) = e^{-(\lambda_1+\lambda_2)t}$ Thus, $Z \sim Expon(\lambda_1 + \lambda_2)$ iii. Probability of Comparing two exponential random variables: Suppose $X \sim Expon(\lambda_1), Y \sim Expon(\lambda_2)$, such that X & Y are independent, then: $p(X > Y) = \frac{\lambda_1}{\lambda_2 + \lambda_1}$ $p(Y > X) = \frac{\lambda_2}{\lambda_2 + \lambda_1}$ iv. Sum of iid exponential random variable follows a gamma distribution: a. Gamma Distribution: Suppose $X \sim \Gamma(n, \lambda)$ $f(x) = \begin{cases} \frac{\lambda e^{-t}(\lambda t)^{n-1}}{(n-1)!}, t \geq 0, \\ 0, \quad t < 0 \end{cases} \qquad E(X) = \frac{n}{\lambda}, \qquad Var(X) = \frac{n}{\lambda^2}$

	b. Let $X_1, X_2, \dots, X_n$ be iid exponential random variables, then $X_1 + X_2 + \cdots + X_n \sim \Gamma(n, \lambda)$
Poisson Process:	Poisson Distribution: suppose $X \sim Pois(\lambda)$: 1. $p(X = k) = \frac{e^{-k}\lambda^k}{k!}$, 2. $E(X) = \lambda$, 3. $Var(X) = \lambda$ Poisson Process: The homogeneous Poisson Process with rate λ is a counting process $N(t)_{t\geq 0}$ satisfying: $N(t) = \max\{n \mid \sum_{i=1}^{n} T_i \leq t\}$, T_i are inter-event time such that $T_i \sim Expon(\lambda)$ i. $N(t) \sim Pois(\lambda t),\ p(N(t) = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, E\big(N(t)\big) = Var\big(N(t)\big) = \lambda t$ Second definition: A counting process $N(t)_{t\geq 0}$ is a rate λ Poisson process if: i. Increments are independent: $N(s) - N(t)\ \&\ N(u) - N(v)$ are independent. ii. $p(N(t+h) - N(t) = 1) = \lambda h + o(h)$ iii. $p(N(t+h) - N(t) \geq 2) = o(h)$ Properties of Poisson Process: 1. Let $N_1(t)_{t\geq 0}$ & $N_2(t)_{t\geq 0}$ be independent Poisson processes with rates λ_1 and λ_2. Let $N(t) = N_1(t) + N_2(t)$, then we know $N(t)_{t\geq 0}$ is a Poisson process with rate $\lambda_1 + \lambda_2$. 2. Conditioned on 1 event happens between $[0, t]$, then the probability of the occurring before any time between the interval $s \in [0, t]$, $p(T_1 < S \mid N(t) = 1) = \frac{s}{t} \sim unif(0, t)$. (Distribution of the conditional first arrival time is uniform). 3. Theorem: Let $N(t)_{t\geq 0}$ be a Poisson process. Conditioned on $N(t) = n$, the n arrival times $S_1, S_2, \dots, S_n$ have the same distribution as the order statistics corresponding to n iid $unif(0, t)$ random variables. $f(S_1, S_2, \dots, S_2 \mid N(t) = n) = \frac{n!}{t^n}$ Poisson Thinning: 1. Homogeneous Poisson Thinning: Let $N(t)_{t\geq 0}$ be a rate λ Poisson process and suppose its events are independently Type 1 with probability p and Type 2 with probability $1-p$. Then, let $N_1(t)$ be the number of type 1 events & $N_2(t)$ be the number of type 2 events, we know $N_1(t)$ & $N_2(t)$ are independent Poisson processes with rates $p\lambda$ & $(1-p)\lambda$ respectively. 2. Non-homogeneous Poisson Thinning: Let $N(t)_{t\geq 0}$ be a rate λ Poisson process & suppose each of its events are classified independently by k types such that the occurrence of event $i \in k$ at time t has the probability of $p_i(t)$. Then, let $N_i(t)$ be the number of occurrences of type i by time t, $N_1(t), \dots, N_k(t)$ are independent Poisson processes with rate $\lambda_i = \lambda \int_0^t p_i(s)ds$.
Continuous Time Markov Chain (CTMT)	Let $\{X(t): t \geq 0\}$ be a collection of random variables, each taking values in N. We say $\{X(t): t \geq 0\}$ is a continuous time MC if for all $s, t \geq 0, i, j \in N$, for all $\{x(u) \in N: u \in [0, s]\}$, the Markov property is satisfied: $p(X(s+t) = j \mid X(s) = i, \{X(u) = x(u)\ for\ all\ u \in [0, s]\}) = p(X(s+t) = j \mid X(s) = i)$ Properties of CTMC: i. Time Homogeneity: $p(X(s+t) = j \mid X(s) = i) = p(X(t) = j \mid X(0) = i)$ ii. Inter-arrival time $T_i \sim Exp(v_i)$ iii. Parameters of CTMC: $p_{ij}, v_i, q_{ij} = p_{ij}v_i, p_{ii} = 0$ iv. Irreducible CTMC: suppose $x, y \in S$, then x, y communicate if $p_{x,y}(t) > 0$ & $p_{y,x}(s) > 0$ for some $s, t \geq 0$. The CTMC is irreducible if there is only 1 communicating class. Chapman Kolmogorov Equations in CTMC: $p_{ij}(s+t) = \sum_{k\in S} p_{kj}(t) * p_{ih}(s)$ Kolmogorov Backwards Equations: $p'_{ij}(t) = \sum_{k\neq i} q_{ik}p_{kj}(t) - v_i p_{ij}(t), p_{ij}(0) = \begin{cases} 0\ if\ i \neq j \\ 1\ if\ i = j \end{cases}$ Kolmogorov Forwards Equation: $p'_{ij}(t) = \sum_{k\neq j} p_{ik}(t)q_{kj} - v_j p_{ij}(t), p_{ij}(0) = \begin{cases} 0\ if\ i \neq j \\ 1\ if\ i = j \end{cases}$ Forming the Matrix Q: $Q_{ij} = \begin{cases} -v_i\ if\ i = j \\ q_{ij} = v_i p_{ij}\ if\ i \neq j \end{cases}$ Solving Differential Equations:

1. Some rules regarding to solving differential equations:
   i. If $f'(x) = a * f(x)$, then $f(x) = C * e^{ax}$, for some constant $C \in R$
2. Procedures of Solving Simple Differential Equations:
   i. Suppose given $f'(x) = a * f(x) + b, f(d) = v \;\; b, a, d, v \in R$, then we write $f(x) = f_1(x) + f_2(x)$
   ii. Then by solving $f_1(x)$ using the rules above, we get: $f_1(x) = C * e^{ax}, C \in R$
   iii. Given that $f_2(x)$ is assumed to be constants, we have $f_2'(x) = 0 = a * f_2(x) + b,\;\; f_2(x) = -\frac{b}{a}$
   iv. Using the $f(d) = v$ to solve for C:

$$C * e^{ad} - \frac{b}{a} = v, \qquad C = \frac{v + \frac{b}{a}}{e^{ad}}$$

Using the balance equation to find the limiting Probabilities: $P = (P_1, \dots, P_n)$
1. Balance Equation: $v_j P_j = \sum_{k \neq j} q_{kj} P_k \,,\; \sum_j P_j = 1$
2. Theorem: For a recurrent, irreducible CTMC,
   i. The limiting probabilities P exist.
   ii. P_i is the long-run proportion of time that CTMC is in state i

Birth & Death Process:

Let $\{X(t), t \geq 0\}$ be a birth & death process with the associated birth rate: $\lambda_i,\; i \in N$ and death rate $\mu_i, i \in N$, then we have the following parameters:
1. $v_i = \mu_i + \lambda_i$ because the time between events $T_i \sim \min\big(Expon(\lambda_i), Expon(\mu_i)\big) \sim Expon(\lambda_i + \mu_i)$
   i. Note: $v_0 = \lambda_0$ because there's no death at 0.
2. $p_{i,i+1} = p\big(Expon(\lambda_i) < Expon(\mu_i)\big) = \frac{\lambda_i}{\mu_i + \lambda_i}$, $q_{i,i+1} = \; v_i * p_{i,i+1} = (\mu_i + \lambda_i)\frac{\lambda_i}{\mu_i + \lambda_i} = \lambda_i$
3. $p_{i,i-1} = p\big(Expon(\mu_i) < Expon(\lambda_i)\big) = \frac{\mu_i}{\mu_i + \lambda_i}$, $q_{i,i-1} = v_i * p_{i,i-1} = \; (\mu_i + \lambda_i) * \frac{\mu_i}{\mu_i + \lambda_i} = \mu_i$

Solving for Limiting Distribution of Birth & Death Process:
Using the balance equation, we have:
1. $v_0 P_0 = q_{1,0} P_1 \rightarrow \lambda_0 P_0 = \mu_1 P_1 \rightarrow P_1 = \frac{\lambda_0}{\mu_1} P_0$
2. $v_1 P_1 = q_{0,1} P_0 + q_{2,1} P_2 \rightarrow (\lambda_1 + \mu_1) P_1 = \lambda_0 P_0 + \mu_2 P_2 \rightarrow (\lambda_1 + \mu_1) P_1 = \mu_1 P_1 + \mu_2 P_2 \rightarrow P_2 = \frac{\lambda_1}{\mu_2} P_1$

   ⋮ ⋮

3. $P_n = \frac{\lambda_{n-1}}{\mu_n} P_{n-1} = \frac{\prod_{i=0}^{n-1} \lambda_i}{\prod_{i=1}^{n} \mu_i} * P_0,\; r_n = \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}}$

Using the fact that $\sum_{i \in n} P_i = 1$, we have:
4. $\sum_{i \in n} P_i = \; \sum_{i=0}^{\infty} r_j P_0 = 1 \rightarrow P_0 = \frac{1}{\sum_{i=0}^{\infty} r_j}$, suppose all λ_i & μ_i are the same, $\sum_{i=0}^{\infty} r_j = \left(\frac{\lambda}{\mu}\right)^n = \frac{1}{1 - \frac{\lambda}{\mu}}$

Two Cases:
   i. If $\sum_{i=0}^{\infty} r_j \rightarrow \infty$, then we do not have the limiting distribution which means that this CTMC is null recurrent or transient.
   ii. If $\sum_{i=0}^{\infty} r_j < \infty$, then the limiting probability exists.

Stationary Distribution of CTMC $P = \{P_1, .. P_n\}$: we can also use the balance equation to find it, or to save time using the detailed equation to get the stationary distribution & reversibility.

Embedded Chain:
Based on the CTMC $\{X(t): t \geq\}$ with transition probabilities p_{ij} & rates v_i, we construct the embedded (discrete time) MC $\{Y_n : n = 0,1, \dots\}$ such that $Y_n = n$th states that $X(t)$ jumps to $= X(S_n)$ where S_n is the time of the nth jump.
1. Properties of the Embedded Chain:
   i. Transition matrix $\tilde{P}$: $\big(\tilde{P}\big)_{i,j} = p_{ij}$
   ii. If the embedded chain is irreducible & finite, then we know that the corresponding CTMC is positive recurrent & irreducible such that it satisfies the limiting distribution theorem above.
   iii. Suppose $\tilde{\pi} = (\pi_1, \dots, \pi_n)$ is the stationary distribution for the embedded chain and let $Z = \; \sum_{i \in n} \frac{\pi_i}{v_i} < \infty$, then the stationary distribution for the original CTMC: $P_i = \frac{1}{z} * \frac{\pi_i}{v_i}$

	Detailed Balance Equation of CTMC: Consider a positive recurrent, irreducible CTMC. If there is a vector $P = (P_0, \ldots P_n)$ satisfying the detailed balance equations: 1. $\sum_{i \in n} P_i = 1$ 2. $P_i * q_{i,j} = P_j * q_{j,i} \quad \forall\, i \neq j, i \,\&\, j \in n$ Then P is the unique stationary distribution & the CTMC is time reversible. Note that if the time reversibility satisfied, then the transition probabilities p_{ij} would be the same for the forward & backward process (same as the discrete case).