CS599

Stochastic Processes

John C.S. Lui

http://www.cse.cuhk.edu.hk/~cslui/csc5420

Stochastic Processes

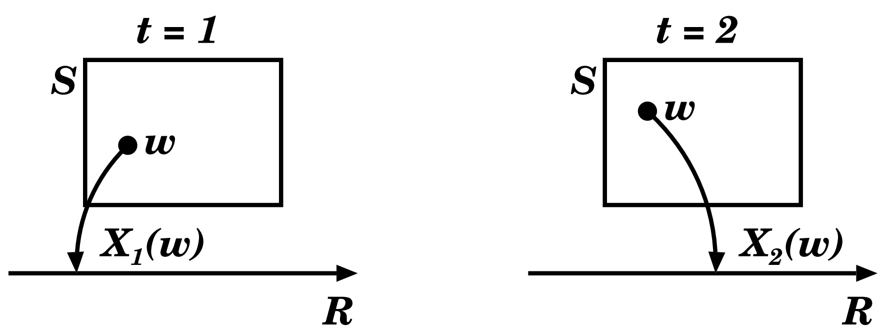
⇒ **Index a "family of r.v.'s" by time** ⇒ **stochastic process**

e.g., $\{ X_t(w) \mid t \in T,\ w \in S \}$ **where** t **is the time index and** s **is the sample space**

⇒ **values assumed by** $X_t(w)$ **are called states of the stochastic process**

⇒ **all possible such values form the state space of the stochastic process**

⇒ **can equivalently denote** $X_t(w) \equiv X(w, t)$



Example

- **Throw a dice three times; the sample space is**

  $$S = \{\ 1, 2, 3, 4, 5, 6\ \}$$

- **Let $X_t(w)$ be defined as follows:**

  $$\forall\ w \in S,\ \ X_1(w) = w,\ \ X_2(w) = 2w,\ \ X_3(w) = 3w$$

- **Then the state space** $= \{\ 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 18\ \}$
- **"State-Time diagram"**

$t = 1$	$t = 2$	$t = 3$
1		
2	2	
3		3
4	4	
5		
6	6	6
	8	
		9
	10	
	12	12
		15
		18

- **We are interested in** $P\,[X_t = 6]$

  $$\Rightarrow\ \ P\,[X_1 = 6,\ \ X_2 = 6,\ \ X_3 = 6]$$
  $$=\ \ P[1^{st}\text{throw} = 6, 2^{nd}\text{throw} = 3, 3^{rd}\text{throw} = 2]$$
  $$=\ \ \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}$$

- **In this example time is discrete and $X_t(w)$ are independent; sometimes will use $X(t),\ \ X_t,\ \ X_n$**

  with understanding that $S = \{w\}$

Characteristics of a Stochastic Process

A. State space (discrete or continuous)
B. The time index (discrete or continuous)
C. Relationship (statistical dependencies) between $\{X_t(w)\}$ (dependence or independence)

$\Rightarrow$ Discrete state space process are called chains
$\Rightarrow$ A discrete time process is often denoted by $X_n, \; n = 0, 1, 2, \cdots$

Distribution of Stochastic Processes

- **At an allowable time t, the PDF of a stochastic process X_t is given by**

$$F_X(x,t) \equiv P\left[X(t) \le x\right]$$

- **For a set of allowable instances, the joint PDF is**

$$F_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) \equiv F_X(\vec{x}; \vec{t})$$

$$\equiv P\left[X_1(t_1) \le x_1,\; X_2(t_2) \le x_2,\; \cdots, X_n(t_n) \le x_n\right]$$

Classification of Stochastic Processes

⇒ **Stationary Process**

- **one where PDF is invariant to shifts in time**

⇒ **for a fixed τ**

$$F_X(\vec{x};\vec{t}) = F_X(\vec{x};\vec{t}+\tau)$$

(i.e., add τ to each element of $\vec{t}$)

⇒ **Independent Process** **(i.e., X_i's are independent r.v.)**

$$F_X(\vec{x};\vec{t}) = F_{X_1}(x_1,t_1)\cdot F_{X_2}(x_2,t_2)\cdot\cdots\cdot F_{X_n}(x_n,t_n)$$

and also $$f_X(\vec{x};\vec{t}) = \prod_{i=1}^{n} f_{X_i}(x_i,t_i)$$ **(continuous state)**

and $$P_X(\vec{x};\vec{t}) = \prod_{i=1}^{n} P_{X_i}(x_i,t_i)$$ **(discrete state)**

⇒ **the dice example is a discrete state, discrete time, independent process; it is not a stationary stochastic process**

Classification of Stochastic Processes (Cont...)

⇒ **Markov Process**

- **allow a restricted form of dependence**

  ⇒ **the future only depends on the current state (doesn't depend on past states or on time spent in the current state or any other prior state)**

  ⇒ **memoryless distribution of time spent in state**

  exponential, geometric

⇒ **discrete state** ⇒ **Markov Chain**

⇒ **for discrete state**

$$P[X(t_{n+1}) = x_{n+1} | X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \cdots, X(t_1) = x_1]$$
$$= P[X(t_{n+1}) = x_{n+1} | X(t_n) = x_n]$$

where $t_1 < t_2 < \cdots < t_n < t_{n+1}$

and x_i is included in some discrete state space

Classification of Stochastic Processes (Cont...)

⇒ **Birth-Death Process**

- **Markovian chains where transitions occur to the nearest neighbors only, i.e., if a process is in state i, the allowable transitions are to i - 1 and i + 1 only**

⇒ **Semi-Markov Process**

⇒ **Markov chain: discrete time ⇒ transition is made at every limit time (Markov property)**

⇒ **means that time spent in each state is geometrically distributed**

Classification of Stochastic Processes (Cont...)

⇒ Relax this restriction, allow the time spent in a state to be arbitrary distributed

or time between state transitions

⇒ semi-Markov discrete time chain

⇒ Note: at time of transition, behaves like an ordinary Markov chain

↳ in these instants we have an embedded Markov chain

⇒ Similarly for continuous-time Markov chains

⇒ transition at any time, but the amount of time spent in a state has an arbitrary distribution or opposed to an exponential distribution

↳ embedded Markov chain is defined at instances of transitions

Classification of Stochastic Processes (Cont...)

⇒ **Random Walks**

- ⇒ **A particle moving among states in some (e.g., discrete) state space**
- ⇒ **Of interest: identifying location of the particle in that space**
- ⇒ **next position = previous position plus r.v. whose value is drawn independent from an arbitrary distribution; this distribution does not change with state of process (except maybe at some boundary states)**
- ⇒ **a sequence of r.v.'s $\{S_n\}$ is referred to as a random walk (starting at the origin) if**

$$S_n = X_1 + X_2 + \cdots + X_n \quad n = 1, 2, \cdots$$

where $S_0 = 0$ and $X_1, X_2, \cdots$ is a sequence of independent r.v.s with a common distribution

Classification of Stochastic Processes (Cont...)

- $\Rightarrow$ **index n counts the number of state transitions the process goes through**
- $\Rightarrow$ **if these constants are taken from discrete set**
  - $\Rightarrow$ **discrete time random walk**
- $\Rightarrow$ **if these constants are taken from continuous set**
  - $\Rightarrow$ **continuous-time random walk**
- $\Rightarrow$ **the interval between these transitions is discrete in an arbitrary way**
  - $\Rightarrow$ **random walk is a special case of a semi-Markov process**
    (often people only care about position after a transition, and so assume meaningless distribution between transitions; then special case of Markov process)

Classification of Stochastic Processes (Cont...)

⇒ if common distribution for X_n, have a discrete-state random walk

⇒ in this case transition probability P_{ij} of going from state i to state j will depend only on the difference in indice j - i (denoted by q_{j-i})

⇒ e.g., of continuous -time random walk

⇒ Brownian motion

e.g., of discrete -time random walk

⇒ total number of heads observed in a sequence of independent coin tosses

Classification of Stochastic Processes (Cont...)

⇒ **Renewal Processes**

⇒ **Count transitions that take place as a function of time**

- **assume $X(0) = 0$ and increases by unity at each transition, i.e., $X(t)$ = number of state transitions made by time t**

⇒ **in this case, a special case of random walk**

$$\text{where } q_1 = 1 \text{ and } q_i = 0, \ \text{where } i \neq 1$$

Classification of Stochastic Processes (Cont...)

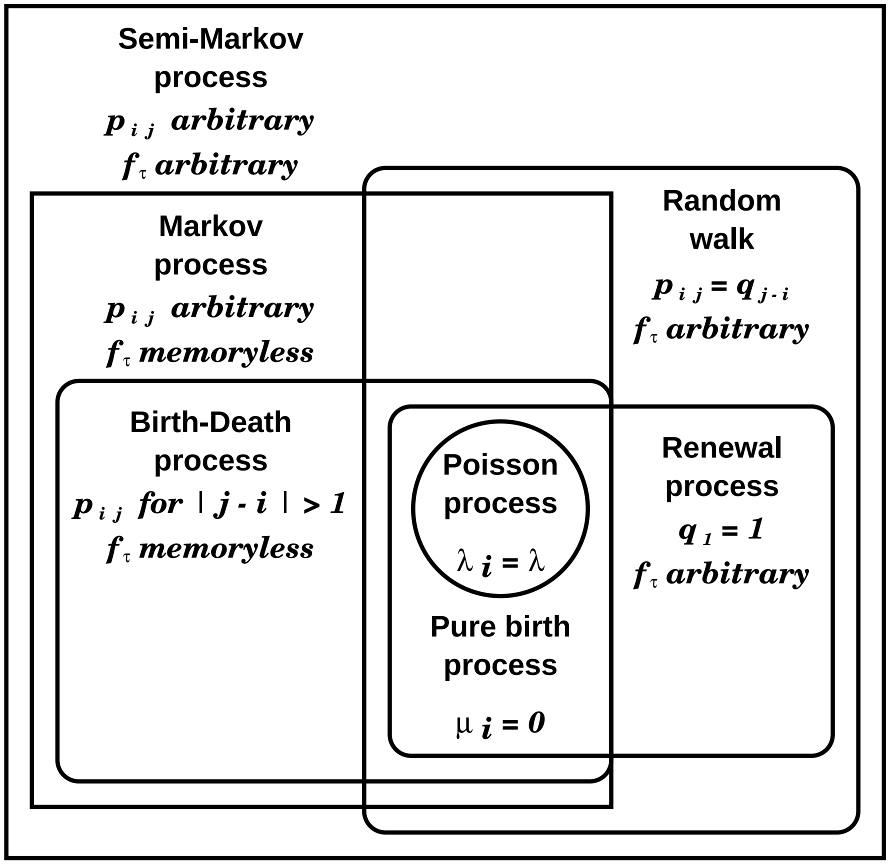
$\Rightarrow$ **can think of:** $S_n = X_1 + X_2 + \cdots + X_n$
as decreasing a renewal process in which S_n **is a r.v. denoting the time at which the** n^{th} **transition takes place**

$\{X_n\}$ **is a set of i.i.d. r.v.s where**
X_n **represents the time between the** $(n - 1)^{th}$ **and** n^{th} **transition**

$\Rightarrow$ **Be careful to distinguish random walk and renewal process. Here above equation describes time of the** i^{th} **renewal transition. Whereas in random walk it describes the state of the process (and the time between transition is some other r.v.)**

Relationships

- **Discrete-State Systems**
  - ⇒ P_{ij} denotes probability of making transition next to state j given the process is in state i
  - ⇒ f_τ denotes distribution of time between transitions (maybe a function of both current and next states of the process)



Discrete Time Markov Chains

⇒ **Let $\{X_n\}$ be a sequence of r.v.'s which assume discrete values**

⇒ **With loss of generality, let n=1, 2, ... correspond to a set of allowable time instants that are obtained from a discrete space**

⇒ **The Markov property can be expended as**

$$P[X_n = j \,|X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \cdots, X_1 = i_1]$$
$$= P[X_n = j \,|X_{n-1} = i_{n-1}] \equiv P_{i_{n-1}j} \quad (***)$$

one step transition probability at step n

⇒ **if transition probabilities are independent of n, then have a homogeneous MC** $\quad (\text{i.e.}, P_{i_{n-1}j} = P_{ij})$

$$P_{ij} \equiv P[X_n = j \,|X_{n-1} = i]$$

do not change with time

(transition probabilities are stationary in time, but this does not have to be a stationary random process)

where $F_X(\vec{x};\vec{t}) = F_X(\vec{x};\vec{t}+\tau)$

(remainder of discussion in terms of homogeneous MCs)

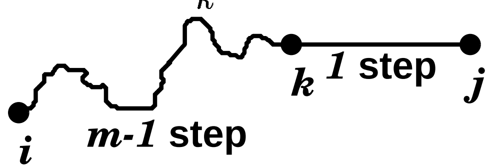
Homogeneous Markov Chains

$\Rightarrow$ **m-step transition probabilities**

- $\Rightarrow$ **probability of various states m steps into the future depends only on m, and not upon current time**
- $\Rightarrow$ $P_{ij}^{(m)} \equiv P\left[X_{n+m} = j \,|X_n = i\right]$

$\Rightarrow$ **from Markov property** $(***)$, **it is easy to know that**

$$P_{ij}^{(m)} = \sum_k P_{ik}^{(m-1)} P_{kj} \qquad m = 2, 3, \cdots$$



- **need to go through some state k**
- **independent, so can multiply probabilities**

Irreducible Markov Chain

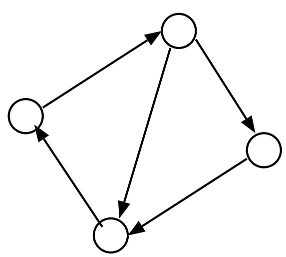
- **A MC is irreducible if every state can be reached from every other state, i.e., if there is m s.t.**

  $$P_{ij}^{(m)} > 0 \text{ and } i, j \in A$$

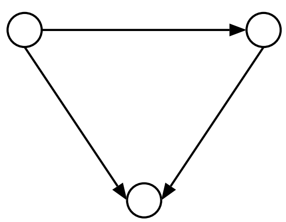
  where A is the set of all states of the MC (all states communicate)

  otherwise, reducible

  E.g.:



Irreducible



Reducible

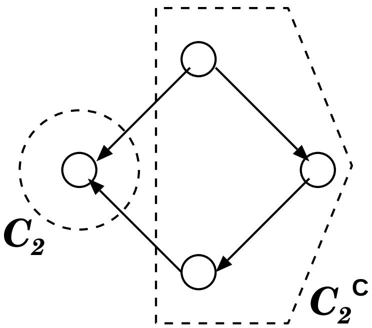
Closed Subset of States

- Let C be a subset of A and C^C be its compliment
- C is a **closed subset** if no one-step transition is possible from any state in C to any state in C^C

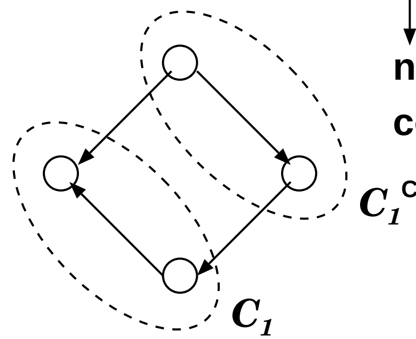
  $\Rightarrow$ if $|C| = 1$, then C is called an **absorbing state**

  necessary and sufficient condition: $P_{ii} = 1$

<u>Ex</u>:



Absorting state



Closed subset

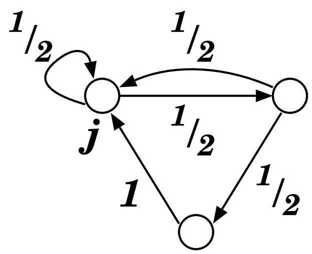
$\Rightarrow$ If C is closed, and it does not include any closed proper subsets of itself, then it is an **irreducible sub-MC**, as defined before

$\Rightarrow$ in above example C_1 is not irreducible, it contains C_2, an absorbing state (closed subset of size 1)

Recurrence

⇒ **Let** $f_i^{(n)} \equiv P[\,\text{the first return to } j \text{ is in } n \text{ steps}\,]$

<u>Ex</u>:



$$f_i^{(1)} = \frac{1}{2} \qquad f_i^{(2)} = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$$

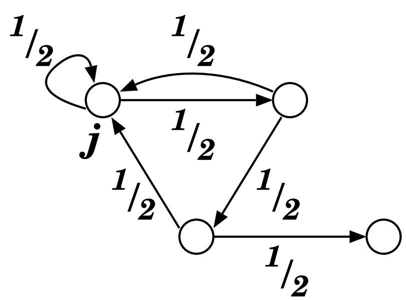
$$f_i^{(3)} = \frac{1}{2}\cdot\frac{1}{2}\cdot 1 = \frac{1}{4} \qquad f_i^{(l)} = 0,\ \text{for } l \leq 4$$

⇒ **Probability of ever returning to state *j* is**

$$f_j = \sum_{n=1}^{\infty} f_j^{(n)}$$

⇒ **in above example, f_j = *1***

<u>Ex</u>:



$$f_i^{(1)} = \frac{1}{2} \qquad f_i^{(2)} = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$$

$$f_i^{(3)} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8} \qquad f_i^{(l)} = 0,\ \text{for } l \leq 4$$

$$\Rightarrow f_j = \sum_{n=1}^{\infty} f_j^{(n)} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + 0 = \frac{7}{8}$$

Resurrence (Cont...)

⇒ Can now classify states of MC according to values of f_j

⇒ Recurrence state $\Rightarrow f_j = 1$

⇒ Transient state $\Rightarrow f_j < 1$

⇒ Mean Recurrence Time

$$M_j \equiv \sum_{n=1}^{\infty} n f_j^{(n)} \text{ for state } j \text{ when } \sum_{n=1}^{\infty} f_j^{(n)} = 1$$

(i.e., for recurrent state)

$\Rightarrow$ if $M_j < \infty$ then j is **recurrent non-null**

$\Rightarrow$ if $M_j = \infty$ then j is **recurrent null**

⇒ Periodicity (for recurrent states)

⇒ if can only return to state j at steps γ, 2γ, 3γ, ... where $\gamma > 1$ and is the largest such integer, then state j is periodic with period γ, otherwise, it is aperiodic

if $\gamma = 1$

State Classification

- **Summary**

States ← of MC

- Recurrent
  - Recurrent Null
    - Periodic
    - Aperiodic
  - Recurrent Non-null
    - Periodic
    - Aperiodic
- Transient

Theorem

⇒ let $\pi_j^{(n)} \equiv P[\, X_n = j \,]$ ⇐ probability of finding the system in state j at n^{th} step

- **Theorem** (without proof)

  The states of an irreducible MC are either all transient or all recurrent non-null or all recurrent null. If periodic, then all states have the same period γ.

⇒ Does there exist a stationary probability distribution $\{\pi_j\}$ describing the probability of bring in state j at some arbitrary time far into the future?

[A probability distribution P_j is said to be a stationary distribution of when we choose it for our initial state distribution, i.e., $\pi_j^{(0)} = P_j$, then for all n we have $\pi_j^{(n)} = P_j$]

⇒ Solving for $\{\pi_j\}$ is a most important part of the analysis of Markov chains

Theorem (Cont...)

⇒ **Next than addresses this**

- **Theorem: In an irreducible and aperiodic homogeneous MC the limiting probabilities**

$$\pi_j \equiv \lim_{n\to\infty} \pi_j^{(n)}$$

always exist and are independent of the initial state probability distribution

Moreover, either

can't be finite !

(a) all states are transient or all states are recurrent null, in which case $\pi_j = 0 \;\forall j$ and there exist no stationary distribution, or

Theorem (Cont...)

(b) all states are recurrent non-null and then $\pi_j > 0$ $\forall j$, in which case the set $\{\pi_j\}$ is a stationary distribution and

$$\pi_j = \frac{1}{M_j}$$

In this case, the quantities π_j are uniquely determined through the following equations

$$1 = \sum_i \pi_i$$
$$\pi_j = \sum_i \pi_i p_{ij}$$

Ergodicity

⇒ **Ergodicity: a state j is ergodic if it is:**

aperiodic, recurrent, and non-null; i.e.,

$$\text{if } f_j = 1, M_j < \infty, \gamma = 1$$

⇒ **if all states of a M.C. are ergodic, the MC is ergodic**

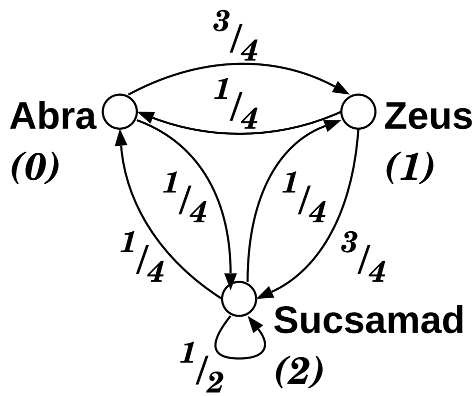
⇒ **a MC is ergodic if the probability distribution $\{\pi_j\}$ as a function of n always converges to a limiting stationary distribution $\{\pi_j\}$, which is independent of the initial state distribution**

⇒ **All state of a finite aperiodic irreducible MC are ergodic**

⇒ **The limiting probabilities of an ergodic MC are often referred to as the equlibrium probabilities (i.e., effect of initial distribution disappeared)**

Example

- **Hippie traveling, waiting to be picked up by car**



State-transition diagram

p_{ij} **permissible direction of road travel**

probability hippie will be picked up by car travel on that road, given he is in current city

$\Rightarrow$ **hippie tries to hitch a ride every day**

remains in same city for another day

$\Rightarrow$ **will refer to #'s on states, *0, 1, 2,* instead now**

$\Rightarrow$ **Transition probability matrix, P, consisting of elements $[\,p_{ij}\,]$**

$\Rightarrow$ **Probability vector** $\vec{\pi} : \vec{\pi} = [\, \pi_0, \pi_1, \pi_2, \cdots \,]$
then we can rewrite the set of equations $\left(\pi_j = \sum_i \pi_i p_{ij} \right)$
as $\vec{\pi} = \vec{\pi} P$

Example (Cont...)

- **In ex:**

$$P = \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

$\Rightarrow$ **Solve:**

$$\left.\begin{array}{lr} \pi_0 = 0 \cdot \pi_0 + \frac{1}{4}\pi_1 + \frac{1}{4}\pi_2 & I \\ \pi_1 = \frac{3}{4}\pi_0 + 0 \cdot \pi_1 + \frac{1}{4}\pi_2 & II \\ \pi_2 = \frac{1}{4}\pi_0 + \frac{3}{4}\pi_1 + \frac{1}{2}\pi_2 & III \end{array}\right\} I = -(II + III)$$

$\Rightarrow$ **linear dependence**

$\Rightarrow$ **Always the case that one equation is linear dependent on others**

$\Rightarrow$ **Need to introduce addition, conservation relationship** $(\text{as is } 1 = \sum_i \pi_i)$ **in order to solve the system**

- **In ex:**

$$1 = \pi_0 + \pi_1 + \pi_2$$

$$\Rightarrow \quad \pi_0 = \frac{1}{5}; \; \pi_1 = \frac{7}{25}; \; \pi_2 = \frac{13}{25}$$

(take any *2* equations and Σ = *1*)

$\Rightarrow$ **equlibrium (stationary) state probability**

Transient Behavior

$\Rightarrow$ **Often interested in transient behavior of system**

$\Rightarrow$ **solving for $\pi_j^{(n)}$ $\Rightarrow$ probability of finding hippie in city j at time n**

$\Rightarrow$ **Define:** $\vec{\pi}^{(n)} \equiv \left[\pi_0^{(n)}, \pi_1^{(n)}, \pi_2^{(n)}, \cdots \right]$

$$\Rightarrow \quad \vec{\pi}^{(1)} = \vec{\pi}^{(0)} P$$

$$\vec{\pi}^{(2)} = \vec{\pi}^{(1)} P = \left[\vec{\pi}^{(0)} P \right] P = \vec{\pi}^{(0)} P^2$$

$$\Rightarrow \quad \vec{\pi}^{(n)} = \vec{\pi}^{(n-1)} P \qquad n = 1, 2, 3, \cdots$$

$$\Rightarrow \quad \vec{\pi}^{(n)} = \vec{\pi}^{(0)} P^n \qquad n = 1, 2, 3, \cdots$$

Recall: $\vec{\pi} = \lim_{n \to \infty} \vec{\pi}^{(n)}$ **assuming the limit exists**

→ previous theorem: if irreducible aperiodic homogeneous MC

$$\Rightarrow \quad \lim_{n \to \infty} \vec{\pi}^{(n)} = \lim_{n \to \infty} \vec{\pi}^{(n-1)} P$$

$$\Rightarrow \quad \vec{\pi} = \vec{\pi} P$$

$\Rightarrow$ **Note: solution for $\vec{\pi}$ is independent of $\vec{\pi}^{(0)}$**

$\Rightarrow$ **HW: try the hippie example with *3* different initial states: *[1, 0, 0], [0, 1, 0], [0, 0, 1]***

Use of Transform in Transient Analysis

$\Rightarrow$ **Define a vector transform:**

$$\vec{\pi}(z) \equiv \sum_{n=0}^{\infty} \vec{\pi}^{(n)} z^n$$

$\Rightarrow$ **apply to** $\vec{\pi}^{(n)} = \vec{\pi}^{(n-1)} P \qquad n = 1, 2, 3, \cdots$

$$\Rightarrow \quad \sum_{n=1}^{\infty} \vec{\pi}^{(n)} z^n = \sum_{n=1}^{\infty} \vec{\pi}^{(n-1)} P z^n$$

$$\Rightarrow \quad \vec{\pi}(z) - \vec{\pi}^{(0)} = z \left(\sum_{n=1}^{\infty} \vec{\pi}^{(n-1)} z^{n-1} \right) P = z\, \vec{\pi}(z)\, P$$

$$\Rightarrow \quad \vec{\pi}(z) = \vec{\pi}^{(0)} \left[I - z P \right]^{-1}$$

$$\Rightarrow \quad \vec{\pi}(z) \Longleftrightarrow \vec{\pi}^{(n)} = \vec{\pi}^{(0)} P^n$$

$$\Rightarrow \quad \left[I - z P \right]^{-1} \Longleftrightarrow P^n$$

P^n **is what we are looking for to get the transient solution**

Example

- **Apply to our ex:**

$$P \;=\; \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \qquad I - zP \;=\; \begin{bmatrix} 1 & -\frac{3}{4}z & -\frac{1}{4}z \\ -\frac{1}{4}z & 1 & -\frac{3}{4}z \\ -\frac{1}{4}z & -\frac{1}{4}z & 1-\frac{1}{2}z \end{bmatrix}$$

$\Rightarrow$ **to invert matrix, must find determinant**

$$\det(I - zP) \;=\; 1 - \frac{1}{2}z - \frac{7}{16}z^2 - \frac{1}{16}z^3 \;=\; (1-z)\;(1+\frac{1}{4}z)^2$$

$$\Rightarrow \quad [\;I - zP\;]^{-1} \;=\; \frac{1}{(1-z)\;(1+\frac{1}{4}z)^2} \times$$

$$\begin{bmatrix} 1-\frac{1}{2}z-\frac{3}{16}z^2 & \frac{3}{4}z-\frac{5}{16}z^2 & \frac{1}{4}z+\frac{9}{16}z^2 \\ \frac{1}{4}z+\frac{1}{16}z^2 & 1-\frac{1}{2}z-\frac{1}{16}z^2 & \frac{3}{4}z+\frac{1}{16}z^2 \\ \frac{1}{4}z+\frac{1}{16}z^2 & \frac{1}{4}z+\frac{3}{16}z^2 & 1-\frac{3}{16}z^2 \end{bmatrix}$$

$\Rightarrow$ **now need inverse transform** $\Rightarrow$ **use partial fraction exp. term by term**

to make it easier, rewrite as sum of 3 matrices, constant, times z, and times z^2

Example (Cont...)

$$\Rightarrow \quad [\ I - zP\ ]^{-1} = \frac{\frac{1}{25}}{1-z}\begin{bmatrix} 5 & 7 & 13 \\ 5 & 7 & 13 \\ 5 & 7 & 13 \end{bmatrix} + \frac{\frac{1}{5}}{(1+\frac{z}{4})^2}\begin{bmatrix} 0 & -8 & 8 \\ 0 & 2 & -2 \\ 0 & 2 & -2 \end{bmatrix}$$

$$+ \frac{\frac{1}{25}}{1+\frac{z}{4}}\begin{bmatrix} 20 & 33 & -53 \\ -5 & 8 & -3 \\ -5 & -17 & 22 \end{bmatrix}$$

$\Rightarrow$ **inverting this:**

$$P^n = \frac{1}{25}\begin{bmatrix} 5 & 7 & 13 \\ 5 & 7 & 13 \\ 5 & 7 & 13 \end{bmatrix} + \frac{1}{5}(n+1)\left(-\frac{1}{4}\right)^n \begin{bmatrix} 0 & -8 & 8 \\ 0 & 2 & -2 \\ 0 & 2 & -2 \end{bmatrix}$$

$$+ \frac{1}{25}\left(-\frac{1}{4}\right)^n \begin{bmatrix} 20 & 33 & -53 \\ -5 & 8 & -3 \\ -5 & -17 & 22 \end{bmatrix}$$

(n = 0, I)
(n = 1, P)

corresponds to equilibrium solution, the other *2* matices decay as $n \to \infty$, corresponds to transient behavior

all rows equal indicates that equilibrium solution is the same regardless of initial state

Remove Homogeneous Assumption

$\Rightarrow$ **DTMC, remove the homogeneous assumption**

$$\text{let } p_{ij}(m,n) \equiv P[\, X_n = j \,|X_m = i \,] \quad n \geq m$$

probability system is in state j in step n, given it was in i at step m

$\Rightarrow$ **must pass through some state q in the middle**

true for all stoch. proc.s

$$\Rightarrow \quad p_{ij}(m,n) \equiv \sum_k P[\, X_n = j, X_q = k \,|X_m = i] \quad \text{for } m \leq q \leq n$$

from def. of cond. prob.

$$\Rightarrow \quad p_{ij}(m,n) \equiv \sum_k P[X_q = k \,|X_m = i]\, P[X_n = j \,|X_m = i, X_q = k]$$

invoke Markov property

$$\Rightarrow \quad P[X_n = j \,|X_m = i, X_q = k] = P[X_n = j \,|X_q = k]$$

$$\Rightarrow \quad p_{ij}(m,n) \equiv \sum_k p_{ik}(m,q)\, p_{kj}(q,n) \quad \text{for } m \leq q \leq n$$

Chapman-Kolmogorov eq. for DTMC

<u>Note</u>: If this was a homogeneous MC, then $p_{ij}(m,n) = p_{ij}^{(n-m)}$
and when n = q + 1, this equation would reduce to our earlier derivation $p_{ij}^{m} = \sum_k p_{ik}^{(m-1)} \; p_{kj}$

Rewrite in Matrix Form

⇒ **Define** $P(n) \equiv [\, p_{ij}(n, n+1) \,]$ **now depends on time**

$P(n) = P$ **if chain is homogeneous**

⇒ **Define** $H(m,n) \equiv [\, p_{ij}(m,n) \,]$ **multistep trans. prob. matrix**

$\Rightarrow H(n, n+1) = P(n)$

⇒ **in the homogeneous case** $H(m, m+n) = P^n$

$\Rightarrow H(m,n) = H(m,q)\, H(q,n) \quad \text{for } m \leq q \leq n$ ← **Chap.-Kol.**

⇒ **require that** $H(n,n) = I$ **(<u>note</u>: all matrices are square ⇒ # states)**

⇒ **since free to choose any *q* in the interval between *m* and *n*: start with *q* = *n* - *1***

$$\Rightarrow p_{ij}(m,n) = \sum_k p_{ik}(m, n-1)\, p_{kj}(n-1, n)$$

$\Rightarrow H(m,n) = H(m, n-1)\, P(n-1)$ ← **forward Chap.-Kol. eq.**

⇒ **also could choose *q* = *m* + *1***

$$\Rightarrow p_{ij}(m,n) = \sum_k p_{ik}(m, m+1)\, p_{kj}(m+1, n)$$

$\Rightarrow H(m,n) = P(m)\, H(m+1, n)$ ← **backward Chap.-Kol. eq.**

Rewrite in Matrix Form

both eq's give same solution

$$\Rightarrow H(m,n) \;=\; P(m)\; P(m+1)\; \cdots\; P(n-1)$$

can check by plugging in above

$\Rightarrow$ **in homogeneous case:** $H(m,n) \;=\; P^{n-m}$

$$\Rightarrow \vec{\pi}(n+1) \;=\; \vec{\pi}(n)\; P(n)$$

solution

$$\Rightarrow \vec{\pi}(n+1) \;=\; \vec{\pi}(0)\; P(0)\; P(1)\; \cdots\; P(n)$$

Continuous-Time Markov Chains

$$P[\ X(t_{n+1}) = j\ \mid X(t_1) = i_1, X(t_2) = i_2, \cdots, X(t_n) = i_n\ \ ]$$
$$= P[\ X(t_{n+1}) = j\ \mid X(t_n) = i_n\ \ ]$$

X(t) $\Rightarrow$ **state of MC at time *t***

$\Rightarrow p_{ij}(s,t) \ \equiv\ P[\ X(t) = j\ \mid X(s) = i\ \ ] \qquad t \geq s$

$\Rightarrow$ **Consider *3* successive time instants** $s \leq u \leq t$

$$\Rightarrow p_{ij}(s,t) \ = \ \sum_k p_{ik}(s,u)\ p_{kj}(u,t)$$

$\Rightarrow$ **Put into matrix form;** $H(s,t) \ \equiv\ [\ p_{ij}(s,t)\ ]$

Chap.-Kal. eq.

$\Rightarrow H(s,t) \ = \ H(s,u)\ H(u,t) \qquad s \leq u \leq t$ **(as before *H(t,t)* = *I*)**

$\Rightarrow$ **Try to derive continuous time analogs of forward and backward equations**

$\Rightarrow$ **Start in forward direction, start with**

$$H(m,n) \ = \ H(m,n-1)\ P(n-1)$$
$$H(m,n) - H(m,n-1) \ = \ H(m,n-1)\ P(n-1) - H(m,n-1)$$
$$H(m,n-1)\ [\ P(n-1) - I\ ] \qquad (*)$$

Continuous-Time Markov Chains (Cont...)

$\Rightarrow$ **Define** $P(t) \equiv [\, p_{ij}(t,\ t+\Delta t)\,]$

$\Rightarrow$ **Let** Δt **be the time step in discrete case**

$\Rightarrow$ **Devide** $(*)$ **by** Δt **and take** $\lim$ **as** $\Delta t \longrightarrow 0$

$\Rightarrow$ $$\frac{\partial\ H(s,t)}{\partial\ t} = H(s,t)\ Q(t) \qquad s \leq t$$

where $$Q(t) = \lim_{\Delta t \longrightarrow 0} \frac{P(t) - I}{\Delta t}$$

infinitesimal generator of $H(s,t)$ or transition rate matrix

$$Q(t) = [\ q_{ij}(t)\ ]$$

define
$\Rightarrow$ $$q_{ii}(t) = \lim_{\Delta t \longrightarrow 0} \frac{p_{ii}(t, t+\Delta t) - 1}{\Delta t}$$

$$q_{ij}(t) = \lim_{\Delta t \longrightarrow 0} \frac{p_{ij}(t, t+\Delta t)}{\Delta t} \qquad i \neq j$$

Continuous-Time Markov Chains (Cont...)

⇒ **Given that we are in state i at time t, probability transition occurs to any other state during interval $(t, t+\Delta t)$ is given by**

$$-q_{ii}(t)\ \Delta t\ +\ o(\Delta t) \qquad\qquad \left(\lim_{\Delta \longrightarrow 0} \frac{o(\Delta t)}{\Delta t}\ =\ 0\right)$$

⇒ **$-q_{ii}(t)$ is the rate at which the process leave state i, when in that state**

⇒ **Similarly the conditional transition probability of going to state j is**

$$q_{ij}(t)\ \Delta t\ +\ o(\Delta t)$$

⇒ **Since** $\displaystyle\sum_j p_{ij}(s,t)\ =\ 1\ \Rightarrow\ \sum_j q_{ij}(t) = 0 \qquad \forall i$

⇒ **Similarly can derive backward Chap.-Kal. eq.**

$$\frac{\partial\ H(s,t)}{\partial\ t}\ =\ -\,Q(s)\ H(s,t) \qquad s \leq t$$

Continuous-Time Markov Chains (Cont...)

$\Rightarrow$ **From forward equation:** **(using individual terms)**

plus some assumption about limits

$$\frac{\partial\ p_{ij}(s,t)}{\partial\ t} = q_{jj}(t)\ p_{ij}(s,t)\ +\ \sum_{k \neq j} q_{kj}(t)\ p_{ik}(s,t)$$

$\Rightarrow$ **Init state i effects the solution through init conditions only:**

$$p_{ij}(s,s) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

$\Rightarrow$ **From backward equation:**

$$\frac{\partial\ p_{ij}(s,t)}{\partial\ s} = -\ q_{ii}(t)\ p_{ij}(s,t)\ -\ \sum_{k \neq i} q_{ik}(s)\ p_{kj}(s,t)$$

where "init" conditions are $p_{ij}(t,t) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

$\Rightarrow$ **Using these equations (unique determine solution):**

$$H(s,t)\ =\ exp\left[\ \int_s^t\ Q(u)\ du\ \right]$$

Satisfies Chap.-Kal. eq.

Analog to discrete case

$$(\text{where } e^{Pt} = I + Pt + P^2\frac{t^2}{2!} + P^3\frac{t^3}{3!} + \cdots)$$

Compute State Probabilities

⇒ **Define** $\pi_j(t) \equiv P[\, X(t) = j \,]$

$$\vec{\pi}(t) \equiv [\, \pi_0(t),\ \pi_1(t),\ \cdots \,]$$

⇒ **Given** $\vec{\pi}(0)$, **can solve for** $\vec{\pi}(t)$

$$\vec{\pi}(t) = \vec{\pi}(0)\ H(0,t)$$

where the general solution is

$$\vec{\pi}(t) = \vec{\pi}(0)\ exp\left[\int_0^t Q(u)\ du\right]$$

Homogeneous Case

$$\Rightarrow p_{ij}(t) \equiv p_{ij}(s, s+t)$$
$$q_{ij} \equiv q_{ij}(t) \quad i, j = 1, 2, \cdots$$
$$H(t) \equiv H(s, s+t) = [\ p_{ij}(t)\ ]$$

$\Rightarrow$ **Chap.-Kal. Eq:** $p_{ij}(s+t) = \sum_k p_{ik}(s)\, p_{kj}(t)$

$\Rightarrow H(s+t) = H(s)\, H(t)$ **(in matrix form)**

$$\frac{d\ H(t)}{d\ t} = H(t)\ Q \qquad \textbf{forward}$$

$$\frac{d\ H(t)}{d\ t} = Q\ H(t) \qquad \textbf{backward}$$

with common initial condition $H(0) = I$

solution

$$\Rightarrow H(t) = e^{Qt}$$

State Probabilties

look at state probability now

⇒ **State probabilities in matrix form**

$$\frac{d\,\vec{\pi}(t)}{d\,t} = \vec{\pi}(t)\,Q$$

⇒ **For an irreducible homogeneous MC, limit exists and independent of initial state of the chain:**

$$\lim_{t \longrightarrow \infty} p_{ij}(t) = \pi_j$$

$\{\pi_j\}$ forms the limiting state probability distribution

⇒ **For an ergodic MC, limit, independent of initial distribution,**

$$\lim_{t \longrightarrow \infty} \pi_j(t) = \pi_j$$

⇒ **This limiting distribution is given uniquely as solution to the following system of linear equations**

$$q_{jj}\,\pi_j + \sum_{k \neq j} q_{kj}\,\pi_k = 0$$

matrix form

⇒ $\vec{\pi}\,Q = 0 \qquad \text{where } \vec{\pi} = [\,\pi_0, \pi_1, \pi_2, \cdots\,]$

⇒ **Compute with** $\sum_j \pi_j = 1,$ **gives us a uniq. sol. to state probs.**

Birth-Death Process

(homogeneous)

- ⇒ **State of system is k (e.g., current population)**
- ⇒ **Birth rate λ_k when population is k**
- ⇒ **Death rate μ_k when population is k**

$$\Rightarrow \lambda_k = q_{k,k+1} \qquad \mu_k = q_{k,k-1}$$

$$(q_{kj} = 0 \text{ for } |k-j| > 1)$$

$$(\text{since} \sum_j q_{kj} = 0, \; qkk = -(\mu_k + \lambda_k)$$

$$Q = \begin{bmatrix} -\lambda_0 & \mu_0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots & \cdots & \cdots & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdots & \cdots & 0 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & 0 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Birth-Death Process (Cont...)

⇒ **Assumptions** needed for B-D process, (in addition to being a homogeneous MC on states $0, 1, 2, \ldots$, that births and deaths are independent (from Markov property),) and

B_1 : P[exactly 1 birth in $(t, t+\Delta t)$ | k in population]

$$= \lambda_k \Delta t + o(\Delta t)$$

D_1 : P[exactly 1 death in $(t, t+\Delta t)$ | k in population]

$$= \mu_k \Delta t + o(\Delta t)$$

B_2 : P[exactly 0 birth in $(t, t+\Delta t)$ | k in population]

$$= 1 - \lambda_k \Delta t + o(\Delta t)$$

D_2 : P[exactly 0 birth in $(t, t+\Delta t)$ | k in population]

$$= 1 - \mu_k \Delta t + o(\Delta t)$$

Solve

⇒ **What** $P_k(t) \equiv P[\,X(t) = k\,] \qquad P_k(t) = \pi_k(t)$

⇒ **Can derive the following from a parallel deviration**

as when did general case

if go through that will get this

$$\frac{d\,P_k(t)}{d\,t} = -(\lambda_k + \mu_k)P_k(t) + \lambda_{k-1}P_{k-1}(t) + \mu_{k+1}P_{k+1}(t) \quad k \geq 1$$

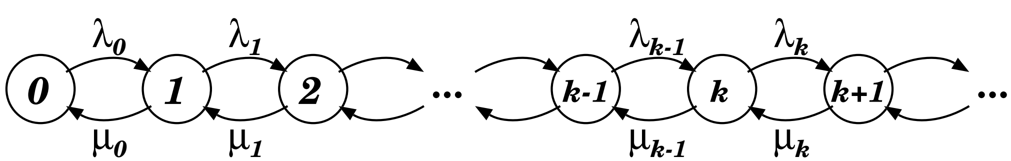
$$\frac{d\,P_0(t)}{d\,t} = -\lambda_0\,P_0(t) + \mu_1\,P_1(t) \quad k = 0$$

set of differential-difference eq's.

$$\text{(to solve need init. conds. and } \sum_{k=0}^{\infty} P_k(t) = 1)$$

⇒ **Try to do the same by "inspection"**

state transition diagram



Solve (Cont...)

⇒ **Rate of change of "flow" into state *k***
= rate of entering *k* - rate of leaving *k*
difference

⇒ **Can derive the following from a parallel deviration**

⇒ **Flow rate into** $k \;=\; \lambda_{k-1}\, P_{k-1}(t) + \mu_{k+1}\, P_{k+1}(t)$

⇒ **Flow rate out of** $k \;=\; (\lambda_k + \mu_k)\, P_k(t)$

⇒ **Difference is the effective prob. flow rate into state k, i.e.,**

flow into a set of states

$$\frac{d\ P_k(t)}{d\ t} = \lambda_{k-1} P_{k-1}(t) + \mu_{k+1} P_{k+1}(t) - (\lambda_k + \mu_k) P_k(t)$$

same as above (haven't talked about boundary cond.)

Pure Birth Process

⇒ **Assume** $\mu_k = 0 \quad \forall k$

⇒ **To simplify, assume** $\lambda_k = \lambda \quad \forall k$

$$\frac{d\ P_k(t)}{d\ t} = -\lambda\ P_k(t) + \lambda\ P_{k-1}(t) \quad k \geq 1 \quad (*)$$

$$\frac{d\ P_0(t)}{d\ t} = -\lambda\ P_0(t) \quad k = 0$$

⇒ **To simplify, assume**

$$P_k(0) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

⇒ **Solving for** $\boldsymbol{P_0(t)}$**, we have**

$$P_0(t) = e^{-\lambda t} \ \Rightarrow \ \text{using in } (*) \text{ for } k = 1$$

$$\Rightarrow \frac{d\ P_1(t)}{d\ t} = -\lambda\ P_1(t) + \lambda\ e^{-\lambda t}$$

sol. $\Rightarrow P_1(t) = \lambda\ t\ e^{-\lambda t}$

Pure Birth Process (Cont...)

⇒ **Continuing by induction**

$$P_k(t) = \frac{(\lambda\ t)^k}{k\ !}\ e^{-\lambda t} \qquad k \geq 0, t \geq 0$$

⇒ **Possion distribution**

pure birth process with constant rate λ

⇒ **given rise to a sequence of birth epochs known as the Poisson Process**

Poisson Process

⇒ **Let k be number of arrivals (from Poisson process) in an interval of length *t***

⇒
$$E[\,K\,] = \sum_{k=1}^{\infty} k\ P_k(t) = e^{-\lambda t} \sum_{k=1}^{\infty} k\ \frac{(\lambda t)^k}{k\ !}$$
$$= e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)\ !} = e^{-\lambda t}\ \lambda\ t \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k\ !}$$

⇒ **Since**
$$e^x = 1 + x + \frac{x^2}{2!} + \cdots$$

⇒ $E[\,K\,] = \lambda t$

⇒ **intuitively, should also see that avg. # of arrivals in *(0,t)* is λt, given that λ is the mean arrival rate**

⇒ **Compute variance:**

$$E[\,K(K-1)\,] = \sum_{k=0}^{\infty} k\ (k-1)\ P_k(t) = e^{-\lambda t} \sum_{k=0}^{\infty} k\ (k-1)\ \frac{(\lambda t)^k}{k\ !}$$
$$= e^{-\lambda t}\ (\lambda t)^2 \sum_{k=2}^{\infty} \frac{(\lambda t)^{k-2}}{(k-2)\ !} = e^{-\lambda t}\ (\lambda t)^2 \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k\ !} = (\lambda t)^2$$

Poisson Process (Cont...)

$\Rightarrow$ $\sigma_k^2 = E[\ K(K-1)\ ] + E[\ K\ ] - (E[\ K\ ])^2 = (\lambda t)^2 + \lambda t - (\lambda t)^2$

$\Rightarrow$ $\sigma_k^2 = \lambda t$

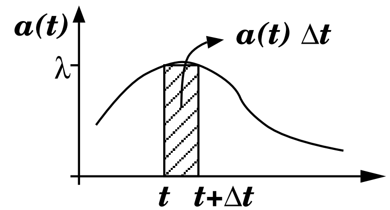
$\Rightarrow$ **Hw: Compute the mean and variance using z-transform**

to get started

$$g_k = P[\ K = k\ ]$$
$$G(z) = E\left[\ z^k\ \right] = \sum_k\ z^k\ g_k$$

Poisson and Exponential Distribution

$\Rightarrow$ **Let r.v. $\tilde{t}$ be time between arrivals with $A(t)$ and $a(t)$, PDF and pdf**



$X \Rightarrow a(t)\ \Delta t + o(\Delta t) \equiv$ **prob. next arrival occur between t and $t+\Delta t$ time sec. units from last arrival**

$$\Rightarrow A(t) = 1 - \underbrace{P\left[\ \tilde{t} > t\ \right]}_{\text{prob. that no arrivals occur in } (0,t) \Rightarrow P_0(t)}$$

$\Rightarrow A(t) = 1 - P_0(t)$

$\Rightarrow$ **In the Poisson case, we have** $A(t) = 1 - e^{-\lambda t} \quad t \geq 0$

$\Rightarrow$ **Differentiate** $\Rightarrow a(t) = \underbrace{\lambda e^{-\lambda t}}_{\text{expo. distri.}} \quad t \geq 0$

$\Rightarrow$ **For a Poisson Process, the time between arrivals is expoenetial distributed**

Poisson and Exponential Distribution (Cont...)

Hw: ① **Show that** $P\left[\, \tilde{t} \leq t + t_0 \,\middle|\, \tilde{t} > t_0 \,\right] = \underbrace{1 - e^{-\lambda t}}$

i.e., cond. distri. is the same as uncond.

② **Compute** $E\left[\, \tilde{t} \,\right]$ **and** $\sigma_{\tilde{t}}^2$ **to show that**

$$E\left[\, \tilde{t} \,\right] = \frac{1}{\lambda} \quad \text{and} \quad \sigma_{\tilde{t}}^2 = \frac{1}{\lambda^2}$$

directly and using Laplace transform