The Exponential Distribution and its Properties

- A widely used distribution in queueing theory is the exponential distribution with parameter $\lambda(>0)$ whose PDF is $f(x) = \lambda e^{-\lambda x}, x \geq 0$. The parameter λ represents a *rate* that has units of events per *time*.
- An exponential RV X satisfies $P\{X > t + s | X > t\} = P\{X > s\}$, for any $t, s \geq 0$. This is called the memoryless property. Further exponential distribution is the only continuous distribution with this property.
- For a continuous positive RV X having distribution function F_X and probability density function f_X, the failure rate or hazard rate function is defined by $r(t) = \frac{f_X(t)}{1-F_X(t)}$. We have that $X \sim Exp(\lambda)$ if and only if $r(t) \equiv \lambda$.
  - $r(t)$ help us to compute the probability that there will be a failure in time $t + dt$ given that there has been no failure up to time t.
- If $X \sim Exp(\lambda)$, $E(X) = 1/\lambda$ and $Var(X) = 1/\lambda^2$. Therefore, the coefficient of variation is 1. $= \frac{\sigma}{\mu}$

- If $X_1, X_2 \ldots, X_n$ are i.i.d. $Exp(\lambda)$, then $\sum_{i=1}^{n} X_i \sim Gamma(n, \lambda)$.
- If $X_1, X_2 \ldots, X_n$ are independent $Exp(\lambda_i)$ then $\min_i X_i \sim Exp(\sum_i \lambda_i)$.
- If $X_1, X_2 \ldots, X_n$ are independent $Exp(\lambda_i)$ then $P\{X_i = \min_j X_j\} = \frac{\lambda_i}{\sum_j \lambda_j}$.
- If $X_1, X_2 \ldots, X_n$ are independent $Exp(\lambda_i)$ with $\lambda_i \neq \lambda_j$ for $i \neq j$ then

$$f_{X_1+X_2+\ldots+X_n}(x) = \sum_{i=1}^{n} C_{i,n} \lambda_i e^{-\lambda_i x},$$

  where $C_{i,n} = \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}$.
- The LT or LST of an exponential RV is $\lambda/(s + \lambda)$.

Counting Processes

- A (continuous-time) random process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ is the number of events occurred from time 0 up to and including time t. For a counting process, we assume
  - (i) $N(0) = 0$.
  - (ii) $N(t) \in \{0, 1, 2, \ldots\}$ for all $t \in [0, \infty)$.
  - (iii) $N(t) - N(s)$ for $0 \leq s < t$ shows the number of events that occur in the interval $(s, t]$.

Examples

$N(t) =$ The number of persons who enter a particular store upto time t — Counting process.

$N(t) =$ Total number of people who were born upto time t — Counting process.

$N(t) =$ the number of persons in a store at a time t — NOT a counting process.

- A counting process $\{N(t), t \geq 0\}$ is said to have independent increments if the number of events that occur in disjoint time intervals are independent. That is, for any $0 \leq t_1 < t_2 < t_3 < \cdots < t_n$, the random variables $N(t_2) - N(t_1), N(t_3) - N(t_2), \ldots, N(t_n) - N(t_{n-1})$ are independent.
- A counting process $\{N(t), t \geq 0\}$ is said to have stationary increments if the distribution of $N(t + s) - N(t)$ depends only on s, for all $s, t \geq 0$.

Poisson Processes (PPs)

- The Poisson process is one of the most widely-used counting processes and usually used in scenarios where we are counting the occurrences of certain events that appear to happen at a certain rate, but completely at random (without a certain structure). For example, the following scenarios could be considered:
  - The number of car accidents at a site or in an area.
  - The location of users in a wireless network.
  - The requests for individual documents on a web server.
  - The outbreak of wars.
  - Photons landing on a photodiode.
  - Arrival of customers at a post office.
- A counting process $\{N(t), t \geq 0\}$ is said to be a (homogeneous) Poisson process (PP) with rate (or intensity) $\lambda > 0$ if
  1. $N(0) = 0$,
  2. it has independent increments, and
  3. the number of events in any interval of length $t > 0$ has $Poi(\lambda t)$ distribution.
- The (homogeneous) Poisson process has stationary increments. The definition above fixes all finite dimensional distributions of the stochastic process.
- Fix any $T > 0$. Define a new process $N_T(\cdot)$ by $N_T(t) = N(T + t) - N(T)$. Then $\{N_T(t)\}$ is again a Poisson process with rate λ. Thus a Poisson process probabilistically restarts itself at any point of time (Markov property).

- Alternative Definition: A counting process $\{N(t), t \geq 0\}$ is said to be a (homogeneous) Poisson process (PP) with rate (or intensity) $\lambda > 0$ if
  1. $N(0) = 0$,
  2. it has independent increments, and
  3. we have the orderliness property:
     $P\{1$ event between t and $t + \Delta t\} = \lambda \Delta t + o(\Delta t)$, and
     $P\{2$ or more events between t and $t + \Delta t\} = o(\Delta t)$,
     where $o(\Delta t)$ denotes a quantity that becomes negligible when compared to Δt as $\Delta t \to 0$.

Exercise

Using the alternative definition and starting from the basic principles, obtain the differential-difference equations satisfied by $p_n(t) = P\{N(t) = n\}$ as
$p_0'(t) = -\lambda p_0(t)$
$p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t), \quad n \geq 1.$
Solve these and show that $N(t) \sim Poi(\lambda t)$ distribution.

Poisson process and Exponential distribution

- If $\{N(t), t \geq 0\}$ is a Poisson process with rate λ, then the times between successive events (also called as inter-event times) are independent and exponentially distributed with rate λ.
  - Let $T_1, T_2, \ldots$ be the inter-event times, where T_n is the time elapsed between $(n-1)$st and nth event.
  - We have $P\{T_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}$ which proves that $T_1 \sim Exp(\lambda)$.
  - For any $s > 0$ and $t > 0$,
    $P\{T_2 > t | T_1 = s\} = P\{$no events in $(s, s+t] | T_1 = s\} = P\{N(t) = 0\} = e^{-\lambda t}$,
    as events in $(s, s+t]$ are not influenced by what happens in $[0, s]$. So T_2 is independent of T_1 and has $Exp(\lambda)$ distribution.
  - Similarly, we can establish that T_3 is independent of T_1 and T_2 with the same distribution, and so on.
- The Poisson process and exponential distribution connection is very important in many ways (for example, in simulating a Poisson process).
- If S_n denotes the time of the nth event, then S_n has $Gamma(n, \lambda)$ distribution.

Example

Consider the car insurance claims reported to insurer. Assume that, the average rate of occurrence of claims is 10 per day. Also, assume that this rate is constant throughout the year and at different times of the day. Further assume that in a sufficiently short time interval of time, there can be at most one claim. What is the probability that there are less than 2 claims reported on a given day? What is the probability that time until the next reported claim is less than 2 hours?

Solution : The given situation can be modelled as Poisson process. The number of arrivals of car insurance claims to the insurer in time intervals of length t is a Poisson process $\{N(t), \ t \geq 0\}$, with rate λ as 10 per day.
Required probability that there are less than 2 claims reported on a given day

$$\begin{aligned} P(N(t+1) - N(t) < 2) &= P(N(1) < 2) \\ &= P(N(1) = 0) + P(N(1) = 1) \\ &= \frac{e^{-\lambda \times 1}(\lambda \times 1)^0}{0!} + \frac{e^{-\lambda \times 1}(\lambda \times 1)^1}{1!} \\ &= e^{-\lambda}(1+\lambda) = \mathbf{11e^{-10}}. \end{aligned}$$

Required probability that time until the next reported claim is less than 2 hours, given at most 2 claims till now (2 hours $= \frac{2}{24}$ days) is $P(T < \frac{2}{24})$, where T is inter-arrival time that follows an exponential distribution with parameter λ. Thus, $P(T < \frac{2}{24}) = 1 - e^{-\lambda \frac{2}{24}} = \mathbf{1 - e^{-\frac{5}{6}}}$.

Poisson Thinning (Splitting or Decomposition) and Superposition

- Consider a Poisson process $\{N(t)\}$ having rate λ. Suppose that each event is classified as a type I event with probability p or a type II event with probability $1-p$, independently of all other events. Let $N_1(t)$ and $N_2(t)$, respectively, denote the number of type I and type II events occurring in $[0, t]$. Note that $N(t) = N_1(t) + N_2(t)$.

  Then, $\{N_1(t)\}$ and $\{N_2(t)\}$ are both Poisson processes with rates λp and $\lambda(1-p)$, respectively. Furthermore, the two processes are independent.
- Let $\{N_1(t)\}$ and $\{N_2(t)\}$ be two independent Poisson processes with rates λ_1 and λ_2, respectively. Then $\{N(t)\}$, where $N(t) = N_1(t) + N_2(t)$, is a Poisson process with rate $\lambda = \lambda_1 + \lambda_2$.

Conditional Distribution of Arrival Times

- Consider a Poisson process $\{N(t)\}$ with rate λ. Given that $N(T) = n$, the n event times $S_1, \ldots, S_n$ have the same distribution as the order statistics corresponding to n independent uniform random variables on $[0, T]$.

> Recall: Let $X_1, X_2, \ldots, X_n$ be IID RVs with PDF $f(\cdot)$. If we let $X_{(i)}$ denote the ith smallest of these RVs, then $X_{(1)}, \ldots, X_{(n)}$ are called the order statistics. The joint PDF of $(X_{(1)}, \ldots, X_{(n)})$ is given by
>
> $$f_{(X_{(1)},\ldots,X_{(n)})}(x_1, \ldots, x_n) = \begin{cases} n! \prod_{i=1}^{n} f(x_i) & \text{for} \quad x_1 < x_2 \ldots < x_n \\ 0 & \text{otherwise}. \end{cases}$$
>
> The marginal PDF of $X_{(i)}$ is
>
> $$f_{X_{(i)}}(x) = \frac{n!}{(n-i)!(i-1)!} f(x)(F(x))^{i-1}(1-F(x))^{n-i}.$$

- Using the above, what we have is that

$$\underline{f_{(S_1,\ldots,S_n)}(s_1, \ldots, s_n)} = \begin{cases} \frac{n!}{T^n}, & \text{for} \quad x_1 < x_2 \ldots < x_n \\ 0 & \text{otherwise}. \end{cases}$$

- The Poisson process is sometimes said to be "completely random". The notion that event times are "completely random" comes from the fac that they are uniformly distributed in time.
- What the uniform property means is that: Given that n events have occurred on $[0, T]$, the *un-ordered* event times are independent and uniformly distributed (or equivalently, the *ordered* event times follow the order statistics of n independent uniform random variables).
- One important consequence of the uniform property of the Poisson process is that the outcomes of random observations of a SP $\{X(t)\}$ have the same probabilities as if the scans were taken at Poisson-selected points.
  ▶ When $\{X(t)\}$ is a queue, this property is called PASTA (Poisson Arrivals See Time Averages).

Compound Poisson Process (CPP)

Let $\{N(t)\}$ be a Poisson process and let $\{Y_i\}$ be an IID sequence of strictly positive integer RVs that are independent of $N(t)$. Then, the SP $\{M(t), t \geq 0\}$, where $M(t) = \sum_{i=1}^{N(t)} Y_i$, is said to be a compound Poisson process (CPP).

▶ If $\{N(t)\}$ is Poisson with rate λ, then $E(M(t)) = \lambda t E(Y_1)$ and $Var(M(t)) = \lambda t E(Y_1^2)$.

Examples

1. Suppose that buses arrive at a sporting event in accordance with a Poisson process, and suppose that the number of fans in each bus are independent and identically distributed. Then $\{M(t), t \geq 0\}$ is a compound Poisson process, where $M(t)$ denotes number of fans who have arrived by time t.

2. Suppose that customers leave a supermarket in accordance with a Poisson process. If Y_i, the amount spent by the ith customer for $i = 1, 2 \ldots$ are IID, then $\{M(t)\}$ is a compound Poisson process, where $M(t)$ denotes the amount of money spent up to time t.

Non-Homogeneous (or nonstationary) Poisson Processes (NHPP)

- An NHPP can be thought of as a PP where λ is replaced by a time-dependent function $\lambda(t)$.
  ⧫ More realistic than a PP (as the rate of occurrence of events are not the same across times).
- A counting process $\{N(t), t \geq 0\}$ is said to be a nonhomogeneous Poisson process (NHPP) (or inhomogeneous or nonstationary Poisson process) with mean event rate $\lambda(t) > 0$ if
  1. $N(0) = 0$,
  2. it has independent increments, and
  3. we have the orderliness property:
     $P\{1 \text{ event between } t \text{ and } t + \Delta t\} = \lambda(t)\Delta t + o(\Delta t)$, and
     $P\{2 \text{ or more events between } t \text{ and } t + \Delta t\} = o(\Delta t)$.
- NHPP loses the stationary increments property.
- For a NHPP $\{N(t), t \geq 0\}$ with mean event rate $\lambda(t)$, the number of events in a time interval $(s, t]$ is a Poisson RV with mean $m(t) - m(s)$, where $m(t) = \int_0^t \lambda(u)du$.
- The function $m(t)$ is sometimes called the mean value function and it represents the cumulative expected number of events by time t.
  [Note: $\lambda(t)$ represents an expected arrival rate].

Example

Consider a nonhomogenous Poisson process with $\lambda(t) = \begin{cases} 5 & \text{if} \quad t \in (1,2], (3,4], \ldots \\ 3 & \text{if} \quad t \in (0,1], (2,3], \ldots \end{cases}$.
Find the probability that the number number of observed occurrences in the time period $(1.25, 3]$ is more than two.

Solution : $N(3) - N(1.25)$ has a Poisson distribution with mean

$$m(3) - m(1.25) = \int_{1.25}^{3} \lambda(t)dt = \int_{1.25}^{2} 5dt + \int_{2}^{3} 3dt = 6.75.$$

Hence, $P\{N(3) - N(1.25) > 2\} = 1 - e^{-6.75}(1 + 6.75 + \frac{(6.75)^2}{2}) = 0.9643.$

Before we wind up PP, note the following.

- The Poisson process is a special case of a larger class of problems called **renewal processes**. A renewal process arises from a sequence of nonnegative IID random variables denoting times between successive events.
- For a Poisson process, the inter-event times are exponential, but for a renewal process, they follow an arbitrary distribution. Many of the properties of the PP also hold true in the renewal context as well (and we will see about the renewal processes later).

Continuous-Time Markov Chains (CTMCs)

- For a CTMC, while the state spaces is discrete, the parameter space is continuous.
  - We will transfer the discrete-time results of MC to CTMC.

S is state space, Xi is any particular state. Ti is the holding time of any particular state. Yt is a stochastic process denoting the state of the process at time t

Definition

Define $S_0 = 0$ and let $\{S_n, n \geq 1\}$ denote a sequence of RVs such that $S_n > S_{n-1}$ for all $n \geq 0$ and $S_n \to \infty$ as $n \to \infty$. Further, let $\{X_n, n \geq 0\}$ be a sequence of RVs taking values in a countable state space S. A stochastic process $\{Y_t, t \geq 0\}$ with $Y_t = X_n$ for $S_n \leq t < S_{n+1}$ is said to be a pure jump process. The variable $T_n = S_{n+1} - S_n$ (resp. X_n) is called the nth **holding time** (resp. the nth **state**) of the process $\{Y_t\}$.

If, further, $\{X_n, n \geq 0\}$ is a Markov chain with (stationary) transition probability matrix $P = ((p_{ij}))$ and the variables T_n are independent and distributed exponentially with parameter λ_{X_n} only depending on the state X_n, then $\{Y_t, t \geq 0\}$ is called a (time-homegeneous) continuous-time Markov chain (CTMC). The chain $\{X_n, n \geq 0\}$ is called the embedded Markov chain of the CTMC.

- We will always assume that $\sup\{\lambda_i, i \in S\} < \infty$.

- A CTMC moves from state to state just like a DTMC, but the time spent in each state is now an exponential random variable.
- Note that if i is not an absorbing state, we can assume that the single-step transition from state i back to itself is not allowed (i.e., $p_{ii} = 0$).
  If i is an absorbing state, then $p_{ii} = 1$ and, in this case, we have $\lambda_i = 0$.
- The process $\{Y_t, t \geq 0\}$ satisfy the Markov property. That is,

  $$P\left\{Y_t = j | Y_{t_n} = i,\, Y_{t_{n-1}} = i_{n-1}, \ldots, Y_{t_0} = i_0\right\} = P\left\{Y_t = j | Y_{t_n} = i\right\}$$

  for all states $i_0,\, i_1,\, \ldots,\, i_{n-1},\, i$ and j in S, for all $n \geq 1$, and for all $t_0, t_1, \ldots, t_n, t$ such that $0 \leq t_0 < t_1 < \cdots < t_n < t$.
- For any $s < t$, define the the **transition probability matrix** $P(s,t)$ from time s to t by its entries $P_{ij}(s,t) = P\{Y_t = j | Y_s = i\}$. And, they satisfy $P_{ij}(s,t) = P_{ij}(0, t-s)$ and we can restrict our attention to $P(t)$ where $P(t) = P(0,t)$.
- With the above notation, the Markov property yields the **Chapman-Kolmogorov equations** $P(s+t) = P(s)P(t)$ for $s, t \geq 0$.
- Note: $P(0) = I$ and the rows of $P(t)$ sum to 1.

Example (Poisson Process)

Define $X_n = n$ for $n = 0, 1, 2, \ldots$. Then $\{X_n, n \geq 0\}$ is a Markov chain with state space $S = \{0, 1, 2, \ldots\}$ and transition probabilities $p_{n,n+1} = 1$ for all $n \geq 0$. Let the holding times T_n be IID exponential with parameter $\lambda > 0$.
The stochastic process $\{Y_t, t \geq 0\}$ with $Y_t = X_n$ for $S_n \leq t < S_{n+1}$ is a CTMC with state space S and is known as Poisson process with rate (or intensity or parameter) λ.

$$\frac{e^{-\lambda} \lambda^k}{k!}$$

Example (Two-state CTMC)

Consider a CTMC with two states $S = \{0, 1\}$ and assume the holding time parameters are given by $\lambda_0 = \lambda_1 = \lambda > 0$.
Since none of the states are absorbing (as $\lambda_i > 0$), and we do not allow self-transitions, the TPM of the embedded MC is $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
We will now determine $P(t)$, starting with $P_{00}(t)$. Assuming $Y_0 = 0$, Y_t will be in 0 if and only if we have an even number of transitions in the time interval $[0, t]$. The time between each transition is an $Exp(\lambda)$ RV and thus the transitions occur according to a Poisson process with parameter λ. We have

$$P_{00}(t) = P\{\text{even number of transitions in}[0,t]\} = \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{2n}}{(2n)!} = \frac{1}{2} + \frac{1}{2}e^{-2\lambda t}.$$

By symmetry, $P_{11}(t) = P_{00}(t)$. We thus have

$$P(t) = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix}.$$

Observe: $\lim_{t \to \infty} P(t) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

Generator Matrix

- A CTMC can be parameterized by the quantities $\{\lambda_i\}$ and $\{p_{ij}\}$. Alternatively, a CTMC can be parameterized by a matrix $Q = ((q_{ij}))$, called **generator matrix** or **infinitesimal generator** or **rate matrix**, and is defined as

$$q_{ij} = \begin{cases} -\lambda_i, & i = j \\ \lambda_i p_{ij}, & i \neq j \end{cases} \quad \text{for all } i, j \in S.$$

- If $Y_0 = i$, the chain will move to the next state at time $T_1 \sim Exp(\lambda_i)$. For small $\Delta t > 0$, $P(T_1 < \Delta t) \approx \lambda_i \Delta t$, i.e., the probability of leaving state i in a short interval of length Δt is approximately $\lambda_i \Delta t$. For this reason, λ_i is often called **the transition rate out of state** i (the expected number of transitions per unit of time).
Formally, we can write $\lambda_i = \lim_{\Delta t \to 0+} \left[\frac{P\{Y_{\Delta t} \neq i | Y_0 = i\}}{\Delta t} \right]$.
- Since the chain moves from state i to state j with probability p_{ij}, we call the quantity $q_{ij} = \lambda_i p_{ij}$, **the transition rate from state** i **to state** j. This is the (i, j)th entry of Q, for $i \neq j$.

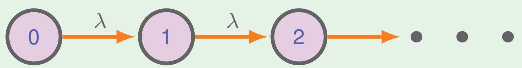
- The diagonal elements (i.e, q_{ii}) of Q are such that the rows of Q sum to 0. That is, $q_{ii} = -\sum_{j \neq i} q_{ij} = -\lambda_i \sum_{j \neq i} p_{ij} = -\lambda_i$ holds for all $i \in S$.
  - If $\lambda_i = 0$, then $\lambda_i \sum_{j \neq i} p_{ij} = \lambda_i = 0$.
  - If $\lambda_i > 0$, then $p_{ii} = 0$ and so $\sum_{j \neq i} p_{ij} = 1$.
- For small $\Delta t > 0$, based on earlier approximation for λ_i, we can obtain $P_{ii}(\Delta t) \approx 1 + q_{ii}\Delta t$ for $i \in S$ and $P_{ij}(\Delta t) \approx q_{ij}\Delta t$ for $i \neq j$.
More precisely, we can state $Q = \lim_{\Delta t \to 0+} \left[\frac{P(\Delta t) - I}{\Delta t} \right]$.
- Q plays a similar role for CTMCs as $P - I$ plays for DTMCs (e.g., stationary equations).
- Q was determined from $\{\lambda_i\}$ and $\{p_{ij}\}$. Alternatively, $\{\lambda_i\}$ and $\{p_{ij}\}$ can be determined from $\{q_{ij}\}$ via $\lambda_i = \sum_{j \neq i} q_{ij}$, $p_{ij} = \frac{q_{ij}}{\sum_{j \neq i} q_{ij}}$.

State Transition Rate Diagram

- The (i, j)th entry of Q is called the **infinitesimal transition rate** from state i to j. A state transition rate diagram (or simply rate diagram) for a CTMC is a directed graph where the nodes represent the states and the edges represent the transition rates q_{ij}. The values of q_{ii} are not shown because they are implied by the other values. As in DTMC, very useful tool here too and we will use extensively.

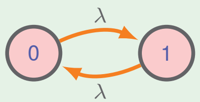
Example (Poisson Process)

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ 0 & -\lambda & \lambda & 0 & \dots \\ 0 & 0 & -\lambda & \lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$



Example (Two-state CTMC)

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}$$



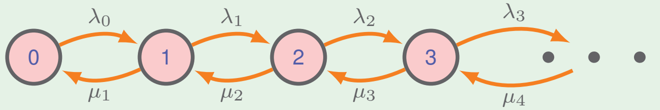
Kolmogorov Forward and Backward Equations

- The transition probabilities $P_{ij}(t)$ of a CTMC satisfy the systems of differential equations
$$\frac{dP_{ij}(t)}{dt} = \sum_{k \in S} P_{ik}(t) q_{kj} \quad \text{and} \quad \frac{dP_{ij}(t)}{dt} = \sum_{k \in S} q_{ik} P_{kj}(t).$$
These are called the Kolmogorov forward and backward equations.respectively.
In matrix notations, $P'(t) = P(t)Q$ and $P'(t) = QP(t)$ (with $P(0) = I$).
- The transition probability matrices can be expressed in terms of the generator by
$P(t) = e^{Qt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n$, for all $t \geq 0$, with Q^n denoting the nth power of Q.
  - Q uniquely determines all transition matrices.
- A CTMC is completely determined (i.e., FDDs are detemined) by the transition matrices and the initial distribution.
- Define $p_i(t) = P\{Y_t = i\}$ for $i \in S$ as the probability that the CTMC is in state i at time t. Denote $\boldsymbol{p}(t)$ as the vector with entries $p_i(t)$ (state probabilities).
- We can characterize the state probabilities via systems of differential equations which are also forms of forward and backward Kolmogorov equations
$$\boldsymbol{p}'(t) = \boldsymbol{p}(t)Q \quad \text{and} \quad \boldsymbol{p}'(t) = Q\boldsymbol{p}(t).$$
  - **Given Q and $\boldsymbol{p}(0)$, we can solve for $\boldsymbol{p}(t)$ from the above.**

Example (Birth-Death Process (BDP))

A *birth-death process (BDP)* is a CTMC $\{Y_t\}$ with $S = \{0, 1, 2, \ldots\}$ in which state transitions either increase the system state by 1 (a birth) or decrease the system state by 1 (a death). The generator matrix (or rate matrix) for a BDP is
$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$
When the system is in state i, births occur with rate λ_i (for $i \geq 0$) and deaths occur with rate μ_i (for $i \geq 1$). In other words, $q_{01} = \lambda_0 = -q_{00}$, and for $i \geq 1$, $q_{i,i+1} = \lambda_i, q_{i,i-1} = \mu_i$ and $q_{ii} = -(\lambda_i + \mu_i)$. Also, $q_{ij} = 0$ for $|i - j| > 1$.

Transition rate diagram of BDP:



The system of differential-difference equations (forward Kolmogorov equations) for the system state probabilities for a BDP are given by
$$p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t)$$
$$p_i'(t) = \lambda_{i-1} p_{i-1}(t) - (\lambda_i + \mu_i) p_i(t) + \mu_{i+1} p_{i+1}(t), \quad i \geq 1.$$

Example (BDP (contd...))

Many queueing systems (where customers arrive/depart one at a time) can be represented as BDPs, where the system state Y_t denotes the number of customers in the system at time t.
- An M/M/1 queue can be modelled by a BDP with $\lambda_i = \lambda$ and $\mu_i = \mu$ for all i.

A BDP is a *pure-birth process* if $\mu_i = 0$ for all i. And, a BDP is a *pure-death process* if $\lambda_i = 0$ for all i.

A Poisson process is also a special case of a BDP with $\lambda_i = \lambda$ and $\mu_i = 0$. Recall that we derived the forward Kolmogorov equations for the system state probabilities from basic principles. It can alternatively be derived from the forward Kolmogorov equations of CTMC by noting that that $q_{ii} = -\lambda, q_{i,i+1} = \lambda$ (for $i \geq 0$), and $q_{ij} = 0$ elsewhere (or equivalently from the forward Kolmogorov equations of the BDP).

Stationary and Limiting Distributions

- As in DTMC, for “nice” CTMCs, a unique stationary distribution exists and equal to the limiting distribution.
- We shall assume the technical assumption $\inf\{\lambda_i, i \in S\} > 0.$
- A CTMC is called **irreducible, transient, recurrent** or **positive recurrent** if the defining Markov chain is.
- Let $\{Y_t\}$ be a CTMC with transition matrix $P(t)$ and state space $S = \{0, 1, 2, \ldots\}$. A probability distribution $\boldsymbol{p}$ on S, i.e, a vector $\boldsymbol{p} = [p_0, p_1, p_2, \ldots]$, where $p_i \in [0, 1]$ and $\sum_{i \in S} p_i = 1$ is said to be a **stationary distribution** for $\{Y_t\}$ if $\boldsymbol{p} = \boldsymbol{p}P(t)$, for all $t \geq 0$.
- The probability distribution $\boldsymbol{p} = [p_0, p_1, p_2, \ldots]$ is called the **limiting distribution** of the CTMC $\{Y_t\}$ if
$$p_j = \lim_{t \to \infty} P\{Y_t = j | Y_0 = i\} \quad \text{for all } i, j \in S, \text{ and we have } \sum_{j \in S} p_j = 1.$$

Example (Two-state CTMC)

Recall the transition matrix, for any $t \geq 0$,

$$P(t) = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{bmatrix}.$$

With $\boldsymbol{p} = [p_0, p_1]$, the equations $\boldsymbol{p} = \boldsymbol{p}P(t), p_0 + p_1 = 1$ gives $p_0 = p_1 = 1/2$ as the stationary distribution. This is also the limiting distribution.
A "nice" chain with a unique stationary distribution that equals the limiting distribution!

- In theory, we can find the stationary (and limiting) distribution by solving $\boldsymbol{p}P(t) = \boldsymbol{p}$, or by finding $\lim\limits_{t\to\infty} P(t)$.
- In practice, finding $P(t)$ itself is usually very difficult. Hence, direct determination of the **steady-state solution** is more difficult.
  - ▶ We need to find alternative ways!

Theorem

For a CTMC, if the embedded DTMC is irreducible and positive recurrent, then there is a unique stationary distribution given by the solution to the stationary equations

$$\mathbf{0} = \boldsymbol{p}Q \quad \text{and} \quad \sum_{i\in S} p_i = 1.$$

Further, under our assumption that the mean holding times in all states are bounded, the chain has a limiting distribution equal to the stationary distribution.

- A CTMC is said to be **regular** it it satisfies the conditions given in the above theorem.
  - ▶ For a regular CTMC, the limit $\lim\limits_{t\to\infty} P_{ij}(t) = p_j$ holds for all $i, j \in S$ and is independent of i.
- Compared to DTMCs, *aperiodicity* is not required for the limiting distribution to exist in a CTMC (as the times between transitions vary continuously). Even if the embedded MC is periodic, the continuous transition times wash out any periodicity that may come from the embedded process.

Balance Equations

- The equation $\mathbf{0} = \boldsymbol{p}Q$ is equivalent to an equation system

$$\sum_{i\neq j} p_i q_{ij} = -p_j q_{jj} \quad \Leftrightarrow \quad \boxed{\sum_{i\neq j} p_i q_{ij} = p_j \sum_{i\neq j} q_{ji}} \quad \text{for all } j \in S.$$

  - ▶ On LHS, $p_i q_{ij}$ is the rate of transitions from i to j (or stochastic flow from i to j in equilibrium). Summing over i gives the overall rate of transitions into state j.
  - ▶ The RHS is the rate of transitions out of state j.
  - ▶ Thus, $\mathbf{0} = \boldsymbol{p}Q$ means that the rate of transitions out of a state equals the rate of transition into the state, in equilibrium (or in steady-state). These are called the (global) **balance equations**.

Example (Poisson Process)

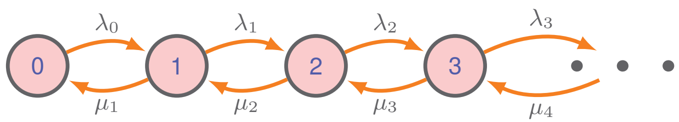
$\boldsymbol{p}Q = \mathbf{0}$ gives $p_0\lambda = 0$ and $p_i\lambda = p_{i-1}\lambda$ for all $i \geq 1$.
This implies that $p_i = 0$ for all $i \in S$ and there is no stationary distribution.

Example (Two-state CTMC)

We have $Q = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}$. Then, $\boldsymbol{p}Q = \mathbf{0}$ gives $p_0 = p_1$, and $p_0 + p_1 = 1$ would then implies that $p_0 = p_1 = 1/2$.

Birth-Death Process (BDP)

- Recall: A CTMC $\{Y_t, t \geq 0\}$ on $S = \{0, 1, 2, \ldots\}$ with the transition rates $q_{i,i+1} = \lambda_i$ for $i \geq 0$, $q_{i,i-1} = \mu_i$ for $i \geq 1$ and $q_{ij} = 0$ for $|i - j| > 1$ is called a **Birth-Death Process (BDP)**.
  - ▶ We assume that $\lambda_i > 0$ for $i \geq 0$ and $\mu_i > 0$ for $i \geq 1$ (and are finite).
- Transition rate diagram of BDP:



- Balance Equations (for the state probabilities):

$$\lambda_0 p_0 = \mu_1 p_1$$
$$(\lambda_i + \mu_i)p_i = \lambda_{i-1}p_{i-1} + \mu_{i+1}p_{i+1}, \quad i \geq 1.$$

- The BDP is irreducible. If it is also positive recurrent, then we will have a unique solution to the above equations and it is called as stationary distribution or limiting distribution or equilibrium distribution or steady-state distribtuion for the BDP.

- We now address the existence of steady-state probability distribution.
- Define two sums:

$$S_1 = \sum_{k=0}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \qquad \text{and} \qquad S_2 = \sum_{k=0}^{\infty} \left(1 \Bigg/ \left(\lambda_k \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) \right)$$

  Case-1: BDP is positive recurrent if and only if $S_1 < \infty$ and $S_2 = \infty$.
  Case-2: BDP is null recurrent if and only if $S_1 = \infty$ and $S_2 = \infty$.
  Case-3: BDP is transient if and only if $S_1 = \infty$ and $S_2 < \infty$.
- It is Case-1 that gives rise to equilibrium probabilities and this is of interest to our studies.
  - Note that the corresponding condition is met whenever the sequence $\{\lambda_n/\mu_n\}$ remains below unity from some n onwards.
- An irreducible BDP on a finite state space $S = \{0, 1, 2, \ldots, N\}$ is said to be a finite-state BDP or finite BDP and is always positive recurrent.