Queueing Theory

Chang-Gun Lee (cglee@snu.ac.kr)

Assistant Professor

The School of Computer Science and Engineering

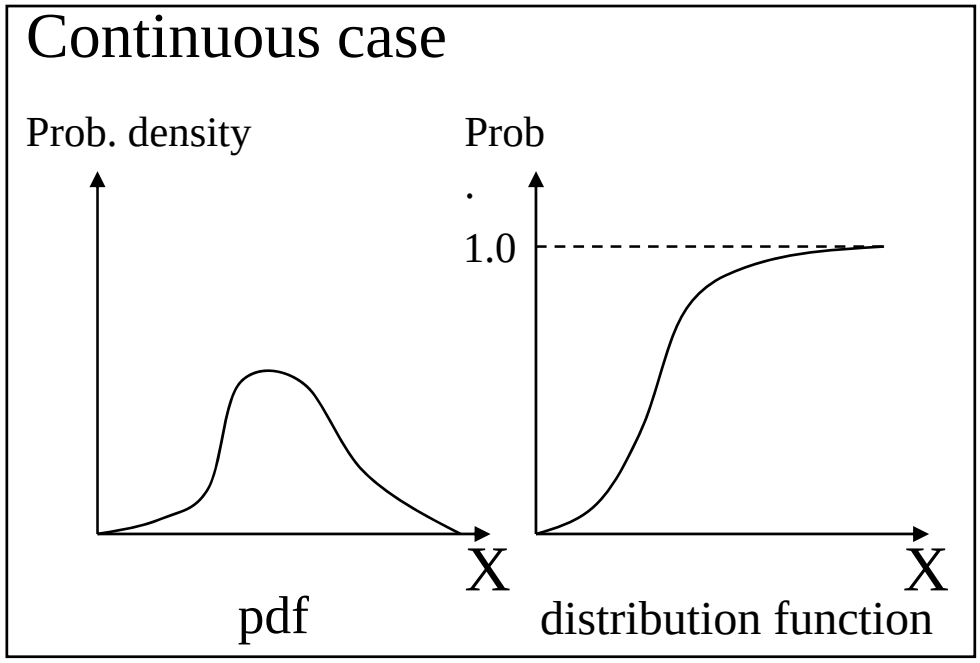
Seoul National University

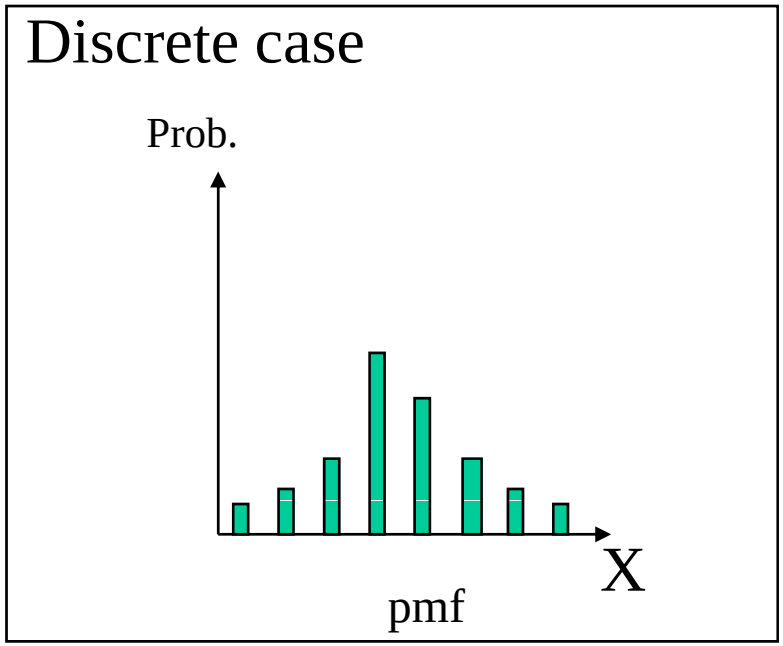
Probability Basics

- Sample space Ω: collection of all possible experimental outcomes
  - E.g., If we look at (sample) a queue, the possible numbers of customers in the queue are {0,1,2,…}
  - E.g., If we sense (sample) a room temperature, the possible outcomes are [-50, +50]
- Sample point ω: one outcome of a sample
  - E.g., The first sample point of the queue = 0;
  - E.g., The first sample point of the temperature = -50 degree
- Event A: set of sample points
  - E.g., Event A: queue is not empty = {1, 2, 3, …}
  - E.g., Event B: the temperature is higher than 10 degree = [+10, +50]
  - Properties
    - P[A] z 0
    - P[Ω] = 1
    - If evnets A and B are mutually exclusive (i.e., $A \cap B = \phi$), then ……. $P(A \cup B) = P(A) + P(B)$

Random Variables

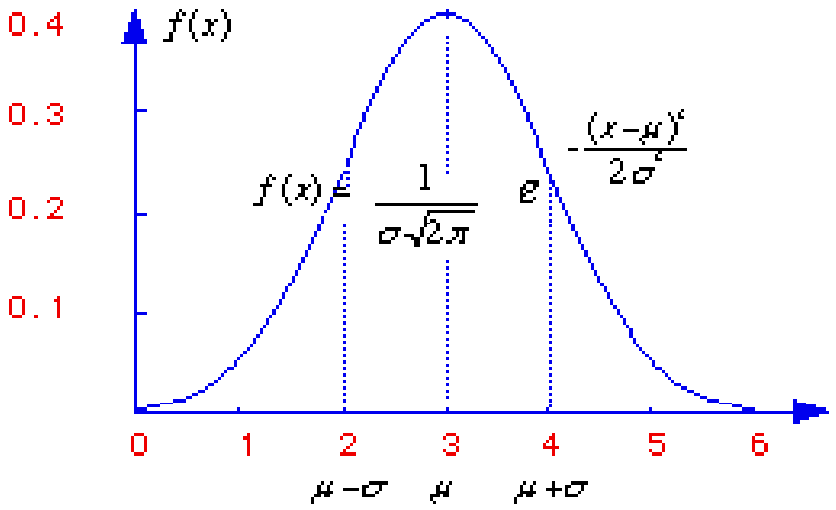
- Random Variable X: a real-valued function that associate a real number with the outcome of an experiment. (a real-valued function defined in the domain of Ω)
  - E.g., X is the modular 10 of customer count in the queue (discrete)
  - E.g., Y is the room temperature in Fahrenheit (continuous)
- A random variable X is usually characterized by probability distribution function (or probability density function, probability mass function)





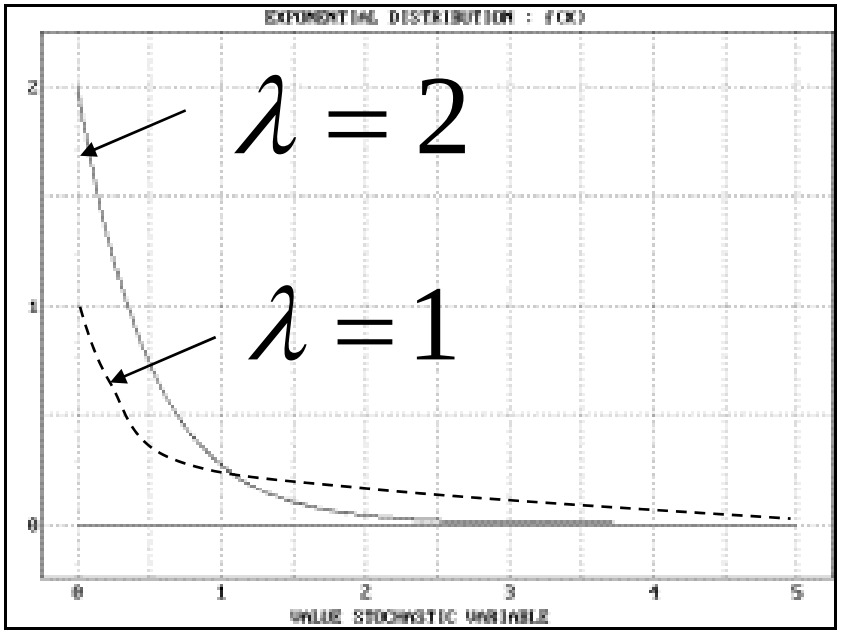
Normal (Gaussian) Distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad \mathrm{E}(X) = \mu, \mathrm{Var}(X) = \sigma^2$$



Exponential Distribution (1)

$$f(x) = \lambda e^{-\lambda x}, x > 0 \qquad \mathrm{E}(X) = \frac{1}{\lambda}, \mathrm{Var}(X) = \frac{1}{\lambda^2}$$



Exponential Distribution (2)

$$f(x) = \lambda e^{-\lambda x}, x > 0 \qquad \mathrm{E}(X) = \frac{1}{\lambda}, \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

Markov property (= memoryless property):

$$P\left[X > t + h \mid X > t\right] = P[X > h]$$

Suppose that X is the wait time for a event. Under the condition that the event does not occur for *t* ($X>t$), the distribution of further waiting is the same as it would be if no waiting time had passed. That is, the system does not remember that *t* time units have produced no event.

$$P\left[X > t + h \mid X > t\right] = \frac{P\left[(X > t + h) \cap (X > t)\right]}{P\left[X > t\right]} = \frac{P\left[(X > t + h)\right]}{P\left[X > t\right]} =$$

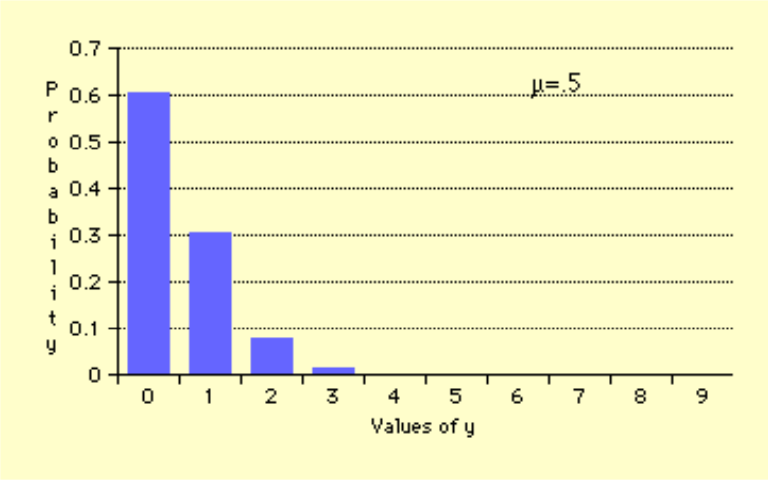
$$\frac{\int_{t+h}^{\infty} \lambda e^{-\lambda x} dx}{\int_{t}^{\infty} \lambda e^{-\lambda x} dx} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda h} = P\left[X > h\right]$$

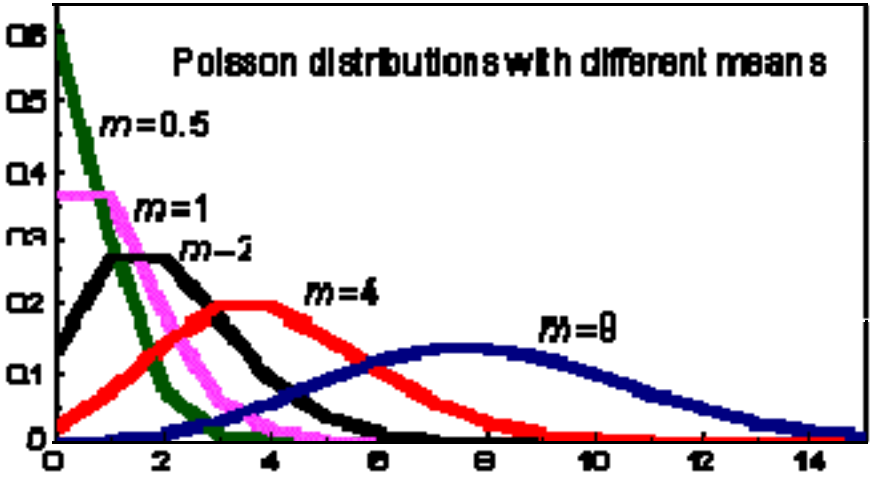
E.g., Inter-arrival time of events with the average arrival rate λ.

E.g., Service time with the average service rate λ.

Poisson Distribution (1)

$$P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!}, x = 0,1,2,\cdots \qquad \mathrm{E}(X) = \lambda, \mathrm{Var}(X) = \lambda$$





Poisson Distribution (2)

- Examples
  - Number of ob arrivals in a unit time
  - Number of jobs completed in a unit time
- Property 1
  - Sum of two Poisson random variables X and Y with the average rate of α and β, respectively, i.e. X+Y, is also a Poisson random variable with rate $\alpha + \beta$.
- Property 2 (Equivalence with Exponential distribution)
  - Consider a Poisson occurrence of events with the rate λ. Let, $0 < t_1 < t_2 < t_3 < \ldots$ be the successive occurrence times of events and let the interarrival times $\{\tau_n\}$ be defined by $\tau_1 = t_1$, $\tau_2 = t_2 - t_1 \ldots \tau_k = t_k - t_{k-1}$. Then the interarrival times $\{\tau_n\}$ are exponential random variables, each with mean $1/\lambda$.
  - Proof

$$P[\tau_n > s] = P[X \cdot s = 0] = \left[e^{-\lambda s} \frac{(\lambda s)^k}{k!} \right]_{k=0} = e^{-\lambda s}$$

$$P[\tau_n \le s] = 1 - e^{-\lambda s}$$

$$f(s) = \lambda e^{-\lambda s}$$

- For modeling a count, use a Poisson distribution. For modeling the time, use an exponential distribution.

Two Important Laws in Probability Theory

- Central Limit Theorem
- Law of large numbers

Central Limit Theorem

- If $X_1, X_2, ..., X_n$ are IID (independent and identically distributed) random varibles with mean μ and variance σ^2, then the sum of the random variables closely follow the normal distribution with mean $n\mu$ and variance $n\sigma^2$ as n → infinity, that is,

$$\sum_{i=1}^{n} X_i \sim \text{normal}(n\mu, \sqrt{n}\sigma) \quad \text{as } n \to \infty$$

Example

- Suppose that on average one job is submitted to a computer system each minute with standard deviation 0.5 and that the numbers of jobs submitted in the sequence of minutes are independent.
- What is the probability that more than 68 jobs are submitted in 64 minutes?

$$X_i : \text{number of jobs submitted in } i\text{-}th \text{ minute}$$

$$X : \text{sum of random variables } X_1, X_2, \cdots X_{64}$$

$$P(X > 68) = 1 - P(X \leq 68) = 1 - P\left(\frac{X - 64}{4} \leq 1\right)$$

$$\approx 1 - P(\text{standardNormalVariable} \leq 1) \approx 0.16$$

Law of Large Numbers

- Weak: If $X_1, X_2, ..., X_n$ are IID (independent and identically distributed) random variables with mean μ, then for all $\varepsilon>0$,

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right| > \varepsilon\right) \to 0 \quad \text{as } n \to \infty$$

- Strong: If $X_1, X_2, ..., X_n$ are independent random variables with $E[X_i] = \mu_i$, then for all ε, $\delta>0$, there exists an integer N such that for all n >= N,

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} \mu_i}{n}\right| < \varepsilon\right) \geq 1 - \delta$$

Example

- In a computer lab, we observed the daily log-on times of 100 users. The average of them is 2hrs 25mins.
- What is the expected log-on time of a user in a day?

$$\frac{\sum_{i=1}^{100} X_i}{100} = 2\text{hr } 25\text{mins}$$

Thus, the expected value of X_i is 2hr 25mins with a high probability

Stochastic process (= Random Process)

- Stochastic process
  - When the probability distribution depends on time, it is better to use a sequence of random variables X(t) depending on time.
  - E.g.: The number of customers in the queue has difference distributions at different times
  - E.g.: The room temperature has different distribution at different times.
  - A stochastic process is a family of random variables, each of which is associated to a time instant t (continuous time parameter *t*, discrete time parameter *n*).

$$\{X(t) : t \in T\} \Rightarrow \text{continuous time process}$$

$$\{X_n : n = 0,1,2,\cdots\} \Rightarrow \text{discrete time process}$$

Two important stochastic process

- Markov Process
- Birth-and-Death Process

Markov Process (1)

- A stochastic process {X(t), t in T} is a Markov process if for any set of times $t_1 < t_2 < \ldots < t_n < t_{n+1}$ in the index set and any set of $\{x_1, x_2, \ldots, x_{n+1}\}$ of n+1 states, we have

$$P\left[X(t_{n+1}) = x_{n+1} \mid X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_n) = x_n\right]$$
$$= P\left[X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n\right]$$

- That is, the future of the process depends only on the present state and not upon the history of the process.
- Markov Chain: if its state space is discrete

Markov Process (2)

Type of time parameter	State Space	
	Discrete	Continuous
Discrete Time	Discrete time Markov Chain	Discrete time Markov Process
Continuous Time	Continuous time Markov Chain	Continuous time Markov Process

Markov Process (3)

- A discrete-time Markov Chain is characterized by the one-step transition probability

$$P\left[X_{n+1} = j \mid X_n = i\right], n, i, j = 0,1,2,\cdots$$

- If the one-step transition probabilities are independent of *n*, such a Markov chain is said to have *Stationary Transition Probability*. ➔ P_{ij}

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Transition Probability Matrix

Example 1

- Consider a sequence of Bernoulli trials in which the probability of success on each trial is p and of failure is q, where $p+q=1$ and $0<p<1$. Let the state of the process at trial n be the number of uninterrupted successes that have been completed at this point. For example, if the first 5 outcomes were SFSSF, we would have $X_0=1$, $X_1=0$, $X_2=1$, $X_3=2$, and $X_4=0$.

- What is the state transition probability matrix?

$$P=\begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}=\begin{bmatrix} q & p & 0 & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & 0 & \cdots \\ q & 0 & 0 & p & 0 & 0 & \cdots \\ q & 0 & 0 & 0 & p & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Example 2

- Consider a communication system that transmits the digits 0 and 1 through several stages. At each stage the probability that the same digit will be received by the next stage, as transmitted, is 0.75. What is the probability that a 0 that is entered at the first stage is received as a 0 by the fifth stage?

$$P = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$$

$$\text{1st stage reception}: \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{pmatrix} 1.0 \\ 0.0 \end{pmatrix}$$

$$\text{2nd stage reception}: \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{pmatrix} 1.0 \\ 0.0 \end{pmatrix}$$

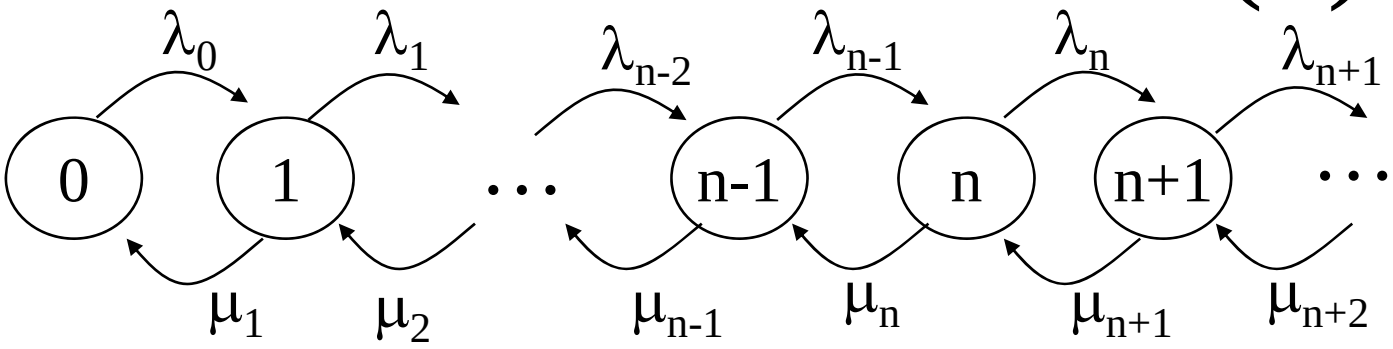
$$\vdots$$

$$\text{5th stage reception}: \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}^5 \begin{pmatrix} 1.0 \\ 0.0 \end{pmatrix} = \begin{bmatrix} 0.515625 & 0.484375 \\ 0.484375 & 0.515625 \end{bmatrix} \begin{pmatrix} 1.0 \\ 0.0 \end{pmatrix}$$

Birth-Death Process (1)

- Consider a continuous time parameter stochastic process $\{X(t), t \geq 0\}$ with the discrete state space 0, 1, 2, ….
- X(t) = n means that the system is in state *n* at time *t*.
- This system is said a birth-death process if there exists non-negative birth rate $\{\lambda_n, n=0,1,2\ldots\}$ and non-negative death rates $\{\mu_n, n=1,2,3,\ldots\}$ such that
  - 1. state changes are allowed only between *n* and *n+1*,
  - 2. the transition probability from n to n+1 between time *t* and *t+h* is $\lambda_n h+o(h)$,
  - 3. the transition probability from n+1 to n between time *t* and *t+h* is $\mu_{n+1} h+o(h)$, and
  - 4. the probability that, in the time interval from *t* to *t+h*, more than one transition occur is $o(h)$.

Birth-Death Process (2)



Let $P\left[X(t)=n\right]$ be $P_n(t)$: Compute $P_n(t)$

$$P_n(t+h) = P_n(t)(1-\lambda_n h-\mu_n h+o(h)) + P_{n-1}(t)(\lambda_{n-1}h+o(h)) + P_{n+1}(t)(\mu_{n+1}h+o(h))$$

$$P_n(t+h) - P_n(t) = -(\lambda_n+\mu_n)hP_n(t) + \lambda_{n-1}hP_{n-1}(t) + \mu_{n+1}hP_{n+1}(t) + o(h)$$

$$\lim_{h\to 0}\frac{P_n(t+h)-P_n(t)}{h} = -(\lambda_n+\mu_n)P_n(t) + \lambda_{n-1}P_{n-1}(t) + \mu_{n+1}P_{n+1}(t)$$

$$\frac{dP_n(t)}{dt} = -(\lambda_n+\mu_n)P_n(t) + \lambda_{n-1}P_{n-1}(t) + \mu_{n+1}P_{n+1}(t)$$

$$\frac{dP_0(t)}{dt} = -\lambda_0 P_0(t) + \mu_1 P_1(t)$$

Birth-Death Process (3)

- In general, finding the time-dependent solutions of a birth-death process is very difficult
- However, if $P_n(t)$ approaches a constant value P_n as $t \rightarrow$ *infinite* for each n, then we say that the system is in "*statistical equilibrium*".

$$\begin{cases} 0 = -(\lambda_n + \mu_n)P_n + \lambda_{n-1}P_{n-1} + \mu_{n+1}P_{n+1} \\ 0 = -\lambda_0 P_0 + \mu_1 P_1 \end{cases}$$

$$\Rightarrow$$

$$\begin{cases} \mu_{n+1}P_{n+1} - \lambda_n P_n = \mu_n P_n - \lambda_{n-1}P_{n-1} \\ \mu_1 P_1 - \lambda_0 P_0 = 0 \end{cases}$$

Let $g_n = \mu_n P_n - \lambda_{n-1}P_{n-1}$. Then

$$\begin{cases} g_{n+1} = g_n \\ g_1 = 0 \end{cases}$$

$$\Rightarrow g_n = 0, n = 1,2,3,... \Rightarrow$$

$$P_{n+1} = \frac{\lambda_n}{\mu_{n+1}} P_n, n = 0,1,2,...$$

$$\Rightarrow$$

$$P_1 = \frac{\lambda_0}{\mu_1} P_0, P_2 = \frac{\lambda_1}{\mu_2} P_1 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} P_0, ...,$$

$$P_n = \frac{\lambda_{n-1}\lambda_{n-2}...\lambda_0}{\mu_n \mu_{n-1}...\mu_1} P_0$$

$$\sum_{n=0}^{\infty} p_n = p_0 + p_1 + p_2 + \cdots = 1, \; p_0 \left(1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} + \cdots \frac{\lambda_{n-1} \cdots \lambda_1 \lambda_0}{\mu_n \cdots \mu_2 \mu_1} + \cdots \right) = 1$$

Birth-Death Process (4)

- The Birth-Death Process is an example of the Continuous time Markov chain.
- A discrete time (h) Birth-Death process is an example of the Discrete time Markov chain. (Characterized by the one-step transition prob.)

$$P\left[X(t+h)=n+1 \mid X(t)=n\right]=\lambda_n h+o(h), \quad n=0,1,2,\cdots$$
$$P\left[X(t+h)=n-1 \mid X(t)=n\right]=\mu_n h+o(h), \quad n=1,2,3\cdots$$
$$P\left[X(t+h)=n \mid X(t)=n\right]=1-\lambda_n h-\mu_n h+o(h), \quad n=0,2,3\cdots$$
$$P\left[X(t+h)=m \mid X(t)=n\right]=0, \text{ for all other cases}$$

- The above probabilities are independent of time parameter *t*. By definition, the discrete time Birth-Death process has "Stationary transition probabilities".
- Supposing h→0 …

$$P=\begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}=\begin{bmatrix} 1-\lambda_0 h & \lambda_0 h & 0 & 0 & \cdots \\ \mu_1 h & 1-\lambda_1 h-\mu_1 h & \lambda_1 h & 0 & \cdots \\ 0 & \mu_2 h & 1-\lambda_2 h-\mu_2 h & \lambda_2 h & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$P_n(t+h)=P_n(t)(1-\lambda_n h-\mu_n h)+P_{n-1}(t)\lambda_{n-1}h+P_{n+1}(t)\mu_{n+1}h$$

- Statistical Equilibrium? → Don't know

Example

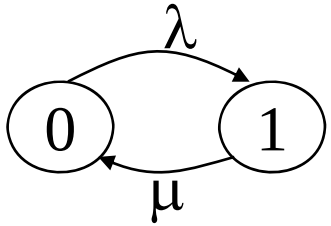
- Consider a computing server with NO waiting line. We assume a Poisson arrival with parameter λ and an exponential service time distribution with parameter μ.
- What is the probability that the server is busy?
- The probability of an arrival in the interval (t, t+h]

$$P[X \geq 1] = 1 - P[X = 0] = 1 - \left(e^{-\lambda h} \frac{(\lambda h)^k}{k!} \right)_{k=0} = 1 - e^{-\lambda h} = \lambda h + o(h)$$

- When the server is busy at *t*, the probability of an service completion by *t*+*h*

$$P[X < h] = 1 - P[X \geq h] = 1 - \int_h^\infty \mu e^{-\mu x} dx = 1 - e^{-\mu h} = \mu h + o(h)$$

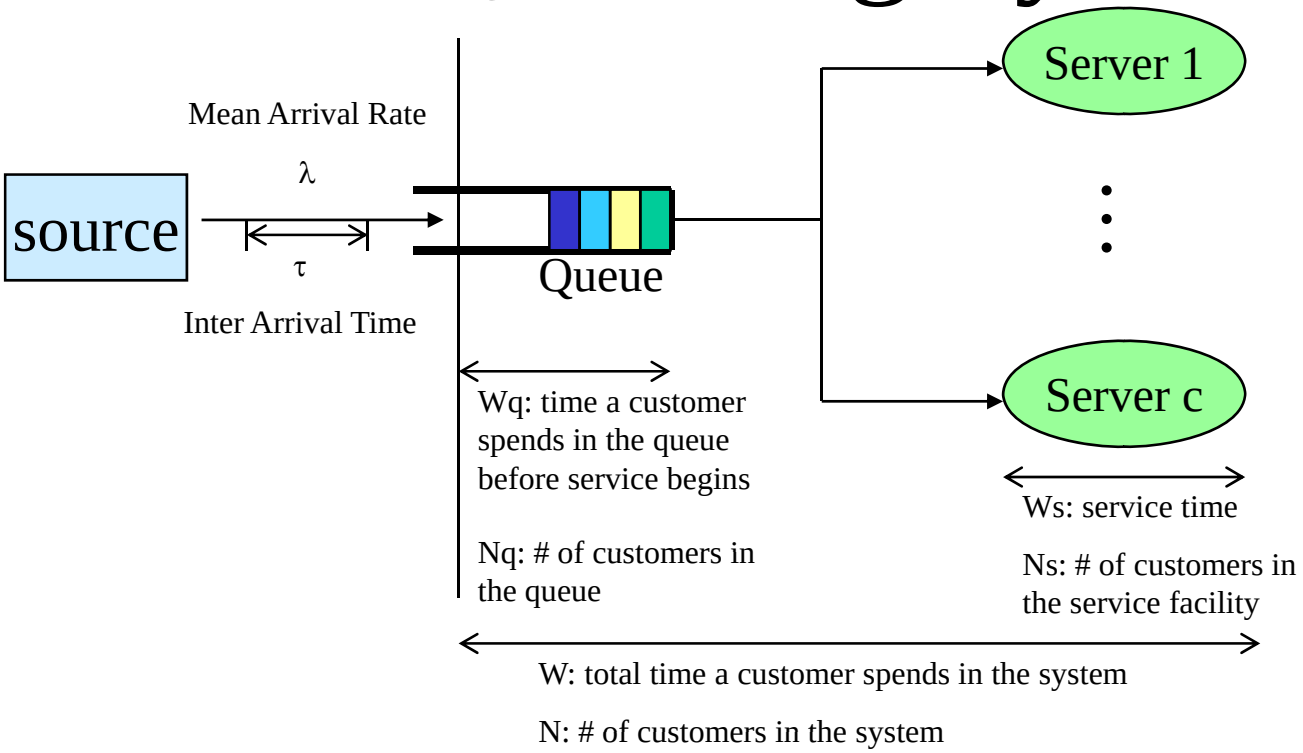
- Birth-death process with $\lambda_0 = \lambda$ ($\lambda_n = 0$ for all other *n*), and $\mu_1 = \mu$ ($\mu_n = 0$ for all other *n*)



$$P_1 = \frac{\lambda_0}{\mu_1} P_0 = \frac{\lambda}{\mu} P_0$$

$$P_0 \left(1 + \frac{\lambda}{\mu} \right) = 1, P_0 = \frac{\mu}{\mu + \lambda}, p_1 = \frac{\lambda}{\mu + \lambda}$$

Queueing System



- Kendall notation: A/B/c/K/m/Z
  - A: inter-arrival time distribution
  - B: service time distribution
  - c: # of servers
  - K: system capacity (queue length + c)
  - m: the number in the population or source
  - Z: queue discipline (FIFO, LCFS, PRI, etc.)

Analysis of Queueing Systems

- The queueing model fits well the birth-death process in general.
- Let *N* be a random variable that represents the number of customers in the system.
- Thus, the steady state probability is ….

$$P_n = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} P_0$$

$$P_0 = \frac{1}{S} \text{ where } S = 1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} + \cdots$$

M/M/1 Queueing

- The number of arrivals follows the Poisson distribution
- The service time follows the Exponential distribution
- The number of servers is ONE

$$\lambda_n = \lambda \text{ for all } n = 0,1,2,\cdots$$

$$\mu_n = \mu \text{ for all } n = 1,2,3,\cdots$$

Let $\frac{\lambda}{\mu} = \rho$

$$S = 1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu^2} + \cdots = 1 + \rho + \rho^2 + \cdots = \frac{1}{1-\rho}$$

Thus,

$$P_0 = 1 - \rho$$

$$P_n = \rho^n (1 - \rho)$$

M/M/1 Queueing

- L=E(N): Expected number of customers in the system

$$L = E(N) = \sum_{n=0}^{\infty} n(1-\rho)\rho^n = \frac{\rho}{1-\rho} \text{ (see geometric dist.)}$$

- VAR(N): Variation of number of customers in the system

$$VAR(N) = \frac{\rho}{(1-\rho)^2} \text{ (see geometric dist.)}$$

- W: Average waiting time of a customer in the system
  - Use Little's Law

Little's Law
(by John D. C. Little)

$$L = \lambda W$$

Avg. # of customers in the system

Customer arrival rate

Avg. time a customer spends in the system

Proof is difficult. See Stidham's paper.

M/M/1 Queueing

- L=E(N): Expected number of customers in the system

$$L = E(N) = \sum_{n=0}^{\infty} n(1-\rho)\rho^n = \frac{\rho}{1-\rho} \text{ (see geometric dist.)}$$

- VAR(N): Variation of number of customers in the system

$$VAR(N) = \frac{\rho}{(1-\rho)^2} \text{ (see geometric dist.)}$$

- W: Average waiting time of a customer in the system
  - Use Little's Law

$$W = \frac{L}{\lambda} = \frac{\frac{\rho}{1-\rho}}{\lambda} = \frac{\rho/\lambda}{1-\rho} = \frac{1/\mu}{1-\rho} = \frac{W_s}{1-\rho}$$

M/M/1 Queueing

- L_q: Expected number of customers in the queue
- W_q: Average waiting time of a customer in the queue

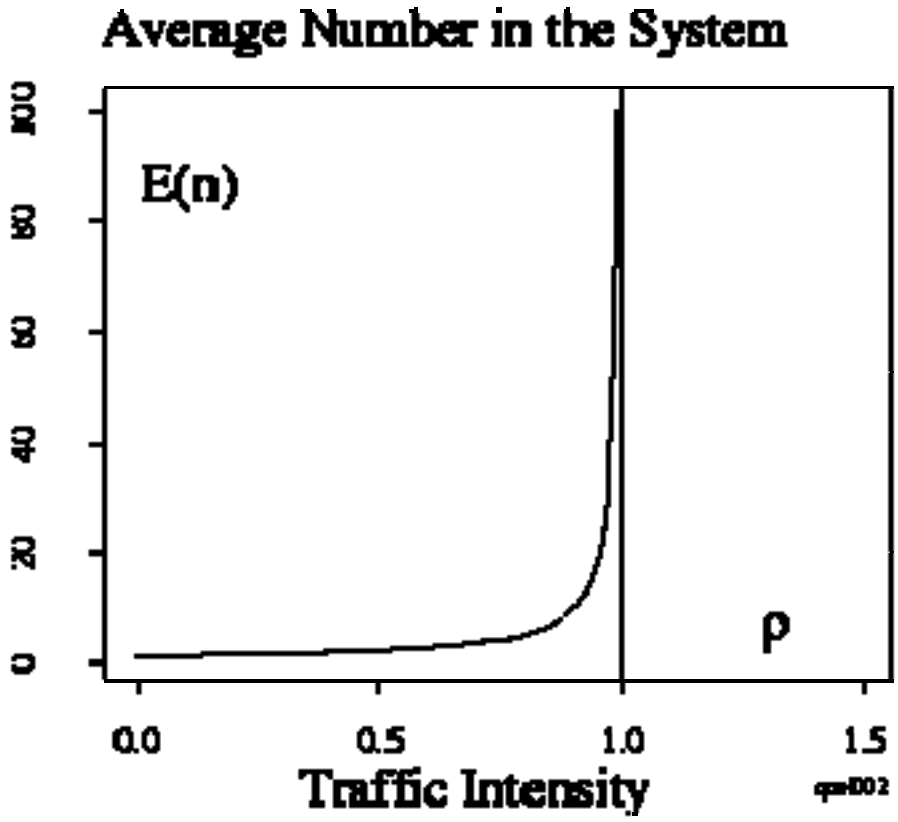
$$W_q = W - W_s = \frac{W_s}{1-\rho} - W_s = \frac{\rho W_s}{1-\rho}$$

$$L_q = \lambda W_q = \lambda \frac{\rho W_s}{1-\rho} = \frac{\rho^2}{1-\rho}$$

- Probability that the server is busy?

$$L_s = \lambda W_s = \lambda \frac{1}{\mu} = \rho$$

M/M/1 Queueing



Example

- Traffic to a message switching center arrives in a random pattern (e.g., exponential distribution) at an average rate of 240 messages per minute. The line has a transmission rate of 800 characters per second. The message length distribution (including control characters) is approximately exponential with an average length of 176 characters.
- L, W, Lq, Wq? $W_s = 176/800 = 0.22\,\text{sec}, \quad \lambda = 240/60\,\text{sec} = 4\,\text{msg/sec}$

  $\rho = \lambda/\mu = 0.88$

  $L = \rho/(1-\rho) = 7.33, \quad W = L/\lambda = 1.83\,\text{sec}$

  $W_q = W - W_s = 1.61\,\text{sec}, \quad L_q = \lambda W_q = 6.44$

- Probability that 10 or more messages are waiting to be transmitted?

$$P(N \geq 11) = \sum_{n=11}^{\infty} (1-\rho)\rho^n = (1-\rho)(\rho^{11} + \rho^{12} + \cdots) = \rho^{11} = 0.245$$

- What would be the average response time W, if the traffic rate into the center increased by 10%?

$$\rho' = (\lambda \times 1.1)/\mu = \rho \times 1.1 = 0.968$$

$$L' = 0.968/(1-0.968) = 30.25, \quad W' = L'/(\lambda \times 1.1) = 30.25/4.4 = 6.875$$

M/M/1/K Queueing

$\lambda_n = \lambda$ for all $n = 0,1,2,\cdots,K-1$

$\mu_n = \mu$ for all $n = 1,2,3,\cdots,K$

Let $\frac{\lambda}{\mu} = \rho$

$$S = 1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu^2} + \cdots + \frac{\lambda^K}{\mu^K} = 1 + \rho + \rho^2 + \cdots + \rho^K = \left(\frac{1-\rho^{K+1}}{1-\rho} \right)$$

Thus,

$$P_0 = \left(\frac{1-\rho}{1-\rho^{K+1}} \right) \quad \text{assuming } \lambda \neq \mu$$

$$P_n = \rho^n P_0 = \rho^n \left(\frac{1-\rho}{1-\rho^{K+1}} \right)$$

What if $\lambda = \mu$?

M/M/1/K Queueing

- L=E(N): Expected number of customers in the system

$$L = E(N) = \sum_{n=0}^{\infty} nP_n = \left(\frac{1-\rho}{1-\rho^{K+1}}\right)\sum_{n=1}^{K}\left(n\rho^n\right) = \left(\frac{1-\rho}{1-\rho^{K+1}}\right)\rho\sum_{n=1}^{K}\left(n\rho^{n-1}\right)$$

$$= \left(\frac{1-\rho}{1-\rho^{K+1}}\right)\rho\sum_{n=1}^{K}\frac{d\rho^n}{d\rho} = \left(\frac{1-\rho}{1-\rho^{K+1}}\right)\rho\frac{d}{d\rho}\sum_{n=1}^{K}\rho^n$$

$$= \left(\frac{1-\rho}{1-\rho^{K+1}}\right)\rho\frac{d}{d\rho}\left(\frac{1-\rho^{K+1}}{1-\rho}\right) = \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}}$$

- W: Average waiting time of a customer in the system
  - Use Little's Law

$\lambda_a = \lambda(1-P_K)$ average arrival rate : arriving is not allowed when $n = K$

$$W = \frac{L}{\lambda_a}$$

Example

- Traffic to a message switching center arrives in a random pattern (e.g., exponential distribution) at an average rate of 240 messages per minute. The line has a transmission rate of 800 characters per second. The message length distribution (including control characters) is approximately exponential with an average length of 176 characters.---same as before. ($\rho = \lambda/\mu = 0.88$)
- If we want to provide only the minimum number of message buffers required to guarantee than Prob(N=K) < 0.005, how many buffers should be provided?

$$P(N = K) = \rho^K \frac{1-\rho}{1-\rho^{K+1}} < 0.005 \Rightarrow K > 25.142607$$

$$\text{no. of buffers} = K - 1 = 25$$

- L, W, Lq, Wq?

$$L = 6.449(cf.7.33), \quad W = 1.62\,\text{sec}(cf.1.83\,\text{sec})$$

$$W_q = W - W_s = 1.40\,\text{sec}(cf.1.61), \quad L_q = 5.573(cf.6.44)$$