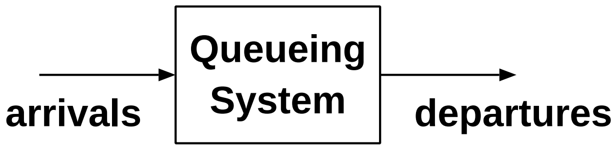
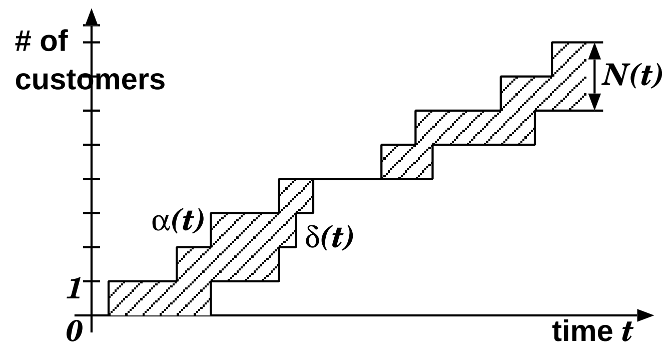
CS599
Birth-Death Process

John C.S. Lui

http://merlot.usc.edu/cs599-s03

Little's Result





⇒ **let** $\alpha(t) \equiv \#$ of arrivals in $(0,t)$

$\delta(t) \equiv \#$ of departures in $(0,t)$

$N(t) = \#$ in system at time t

$N(t) = \alpha(t) - \delta(t)$

⇒ **total area between these two curves up to *t* represents the total time all customers spent in the system (in units of customer-seconds) during interval *(0,t)***

⇒ **denote by** $\gamma(t)$

⇒ **let** λ_t **be avg. arr. rate (cust/sec) during *(0,t)***

$$\Rightarrow \lambda_t \equiv \frac{\alpha(t)}{t}$$

Little's Result (Cont...)

$\Rightarrow$ ***T_t*** **= system time per cust. avg. over all customers in *(0,t)***

$$\Rightarrow T_t = \frac{\gamma(t)}{\alpha(t)}$$

$\Rightarrow$ **let $\overline{N_t}$ = avg. # of cust. in system during *(0,t)***

$$\Rightarrow \overline{N_t} = \frac{\gamma(t)}{t}$$

$$\Rightarrow \overline{N_t} = \lambda_t T_t$$

Assume $\lambda = \lim_{t\to\infty} \lambda_t$ **and** $T = \lim_{t\to\infty} T_t$ **exist** $\Rightarrow \overline{N} = \lim_{t\to\infty} \overline{N_t}$ **exists**

$\Rightarrow \overline{N} = \lambda T$ **Little's Result** $(\overline{N_q} = \lambda W)$

General Equilibrium Solution

(Birth-Death Process)

⇒ **Assume** $p_k \equiv \lim_{t \to \infty} P_k(t)$ **exist (prob. of finding birth-death system in state** $\boldsymbol{k}$**)**

↳ **equil. prob. of finding k cust's in the system**

⇒ **The change in input and output flow at the** $\boldsymbol{lim}$ **is** $\boldsymbol{0}$ **or input flow = output flow** ⇒ **conservation of flow**

Flow rate into state $\boldsymbol{k}$ $= \lambda_{k-1} p_{k-1} + \mu_{k+1} p_{k+1}$

Flow rate out of state $\boldsymbol{k}$ $= (\lambda_k + \mu_k) p_k$

In equil.

⇒ $\lambda_{k-1} p_{k-1} + \mu_{k+1} p_{k+1} = (\lambda_k + \mu_k) p_k$

⇒ **of course,** $\sum_k p_k = 1$

⇒ **Can apply conservation of flow to any subset of states (not just one state)** (subset of states: **choose boundary**)

⇒ **If counter flow into/from subset** $\boldsymbol{0}$ **through** $\boldsymbol{k\text{-}1}$**:**

$$\lambda_{k-1} p_{k-1} = \mu_k p_k$$

General Equilibrium Solution (Cont...)

$$\Rightarrow p_1 = \frac{\lambda_0}{\mu_1} p_0$$

$$\Rightarrow 0 = -(\lambda_1 + \mu_1) p_1 + \lambda_0 p_0 + \mu_2 p_2$$

$$0 = -(\lambda_1 + \mu_1) \frac{\lambda_0}{\mu_1} p_0 + \lambda_0 p_0 + \mu_2 p_2$$

$$0 = -\frac{\lambda_1 \lambda_0}{\mu_1} p_0 - \lambda_0 p_0 + \lambda_0 p_0 + \mu_2 p_2$$

$$\Rightarrow p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0$$

$\Rightarrow$ **From above can guess that general solution is**

$$p_k = \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} p_0$$

(can verify with p_{k+1})

$$\Rightarrow p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \qquad k = 0, 1, 2, \cdots$$

$\Rightarrow$ **from** $\sum_k p_k = 1 \;\Rightarrow\; p_0 = \dfrac{1}{1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}}$

Existence of Steady-State Probabilities

⇒ **Existence** of steady-state probabilties p_k

⇒ For above expression to represent a probability distribution, **usually** require $p_0 > 0$ → require that system empties occationally

⇒ More formally/precisely

Define:

$$S_1 \equiv \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}$$

$$S_2 \equiv \sum_{k=1}^{\infty} \left(\frac{1}{\lambda_k \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}} \right)$$

⇒ All states:

- **ergodic** iff $S_1 < \infty$, $S_2 = \infty$ ⇒ equil. prob's
- **recurrent null** iff $S_1 = \infty$, $S_2 = \infty$
- **transient** iff $S_1 = \infty$, $S_2 < \infty$

Existence of Steady-State Probabilities (Cont...)

$\Rightarrow$ **Conditon for ergodicity is met whenever seqence** $\left\{\frac{\lambda_k}{\mu_k}\right\}$ **remains below unity for some k onwards, i.e.,**

there exists some k_0 s.t. $\forall k \geq k_0 \quad \frac{\lambda_k}{\mu_k} < 1$

M/M/1 System

- $\lambda_k = \lambda \qquad k = 0, 1, 2, \cdots$ **(inf. queueing space)**

  $\mu_k = \mu \qquad k = 0, 1, 2, \cdots$

  $$\Rightarrow p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^k \qquad k \geq 0$$

  $\Rightarrow$ **To be ergodic (and hence $p_k > 0$)** $\Rightarrow S_1 < \infty$ and $S_2 = \infty$

  $$\Rightarrow S_1 = \sum_{k=0}^{\infty} \frac{p_k}{p_0} = \underbrace{\sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k}_{\textbf{converges iff } \frac{\lambda}{\mu} < 1} < \infty$$

  $$\Rightarrow S_2 = \sum_{k=0}^{\infty} \frac{1}{\lambda \left(\frac{p_k}{p_0}\right)} = \underbrace{\sum_{k=0}^{\infty} \frac{1}{\lambda} \left(\frac{\mu}{\lambda}\right)^k = \infty}_{\textbf{satisfied if } \frac{\lambda}{\mu} \leq 1}$$

  $\Rightarrow$ **necessary and sufficient cond. for ergodic is $\lambda < \mu$**

M/M/1 System

$$\Rightarrow p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^k}$$

(converges since λ < μ)

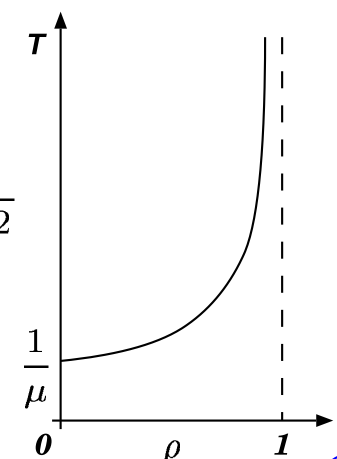
$$\Rightarrow p_0 = \frac{1}{1 + \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}}} = 1 - \frac{\lambda}{\mu}$$

$\Rightarrow$ **let** $\rho = \frac{\lambda}{\mu} \Rightarrow$ **util.** $\Rightarrow 0 \leq \rho < 1$ **(for stability)**

→ **examines** $p_0 > 0$

$$\Rightarrow p_k = (1-\rho)\rho^k \qquad k = 0, 1, 2, \cdots$$

← **depends only on ratio** $\frac{\lambda}{\mu}$



Hw: rewrite $\overline{N} = \sum_{k=0}^{\infty} k p_k = \frac{\rho}{1-\rho}, \quad \sigma_N^2 = \frac{\rho}{(1-\rho)^2}$

$$T = \frac{\frac{1}{\mu}}{1-\rho}$$

(Little's Result)

M/M/1 System (Cont...)

utilization factor

⇒ ρ = **Ratio of the rate at which work enters the system and the max rate at which the system can perform this work**

⇒ **Single server**

$$\rho \equiv \text{ evg. arr. rate of customers} \times \text{avg. service time} = \lambda \overline{X}$$

amount of work each customer ings

⇒ **can do this work at rate of *1* sec./sec.**

⇒ **avg: λ cust. arr. per sec.**

⇒ **Multiple servers**

now work capacity of system is m sec./sec. then

$$\Rightarrow \rho \equiv \frac{\lambda \overline{X}}{m}$$

(Note: above time of service rate is indep. of system state)

⇒ **As long as** $0 \leq \rho < 1$ **can interpret it as**

$$\rho = E[\text{ fraction of busy servers }]$$

M/M/m System

(m server case)

$\Rightarrow \lambda_k = \lambda \qquad k = 0, 1, 2, \cdots$

$\Rightarrow \mu_k = \min[\, k\mu, m\mu \,] = \begin{cases} k\mu & 0 \le k \le m \\ m\mu & m \le k \end{cases}$

$\Rightarrow$ **Condition for ergodicity is** $\dfrac{\lambda}{m\mu} < 1$

$\Rightarrow$ **for** $k \le m$:

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}$$

$\Rightarrow$ **for** $k \ge m$:

$$p_k = p_0 \prod_{i=0}^{m-1} \frac{\lambda}{(i+1)\mu} \prod_{j=m}^{k-1} \frac{\lambda}{m\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{m!m^{k-m}}$$

$\Rightarrow p_k = \begin{cases} p_0 \frac{(m\rho)^k}{k!} & k \le m \\ p_0 \frac{\rho^k m^m}{m!} & k \ge m \end{cases}$ **where** $\rho = \dfrac{\lambda}{m\mu} < 1$ ← **expected fraction of busy servers**

M/M/m System (Cont...)

⇒ **Solve for p_0:**

$$p_0 = \left[1 + \sum_{k=1}^{m-1} \frac{(m\rho)^k}{k!} + \sum_{k=m}^{\infty} \frac{(m\rho)^k}{m!} \frac{1}{m^{k-m}} \right]^{-1}$$

$$= \left[\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \left(\frac{(m\rho)^m}{m!} \right) \left(\frac{1}{1-\rho} \right) \right]^{-1}$$

$$p[\text{ queueing }] = \sum_{k=m}^{\infty} p_0 \frac{(m\rho)^k}{k!} \frac{1}{m^{k-m}}$$

↓ **arr. cust. joins queue**

$$= \frac{\left(\frac{(m\rho)^m}{m!} \right) \left(\frac{1}{1-\rho} \right)}{\left[\sum_{k=1}^{m-1} \frac{(m\rho)^k}{k!} + \left(\frac{(m\rho)^m}{m!} \right) \left(\frac{1}{1-\rho} \right) \right]}$$

⇐ **Erlang's C formula**

probability that no telephone trunk is available for an arriving call

M/M/m/m System

(m-server loss system)

$$\Rightarrow \lambda_k = \begin{cases} \lambda & k < m \\ 0 & k \geq m \end{cases}$$

$$\mu_k = k\mu \qquad k = 1, 2, \cdots, m$$

$$\Rightarrow p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} \qquad k \leq m$$

$$p_k = \begin{cases} p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} & k \leq m \\ 0 & k > m \end{cases}$$

$\Rightarrow$ **Solve for p_0:**

$$p_0 = \left[\sum_{k=0}^{m} \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} \right]^{-1}$$

$\Rightarrow$ **p_m: fraction of time that all servers are busy**

$$p_m = \frac{\left(\frac{\lambda}{\mu}\right)^m \frac{1}{m!}}{\sum_{k=0}^{m} \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}}$$

Erlang's loss formula
or
Erlang's B formula

Homework

HW: Derive p_k, p_0 for:

- M/M/1//M ⇒ finite cust. population
- M/M/∞//M ⇒ and inf. server
- M/M/m/K/M ⇒ finite pop., m-servers, finite storage (K)

assume $M > K > m$