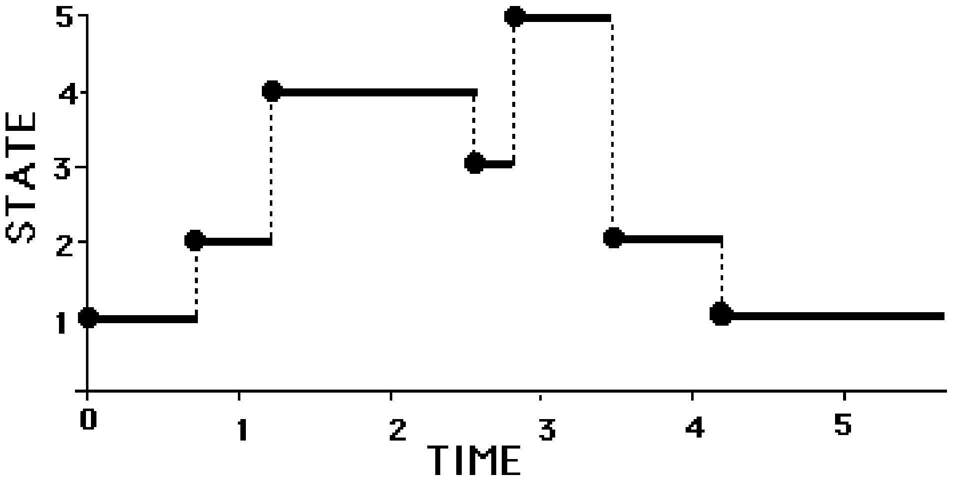
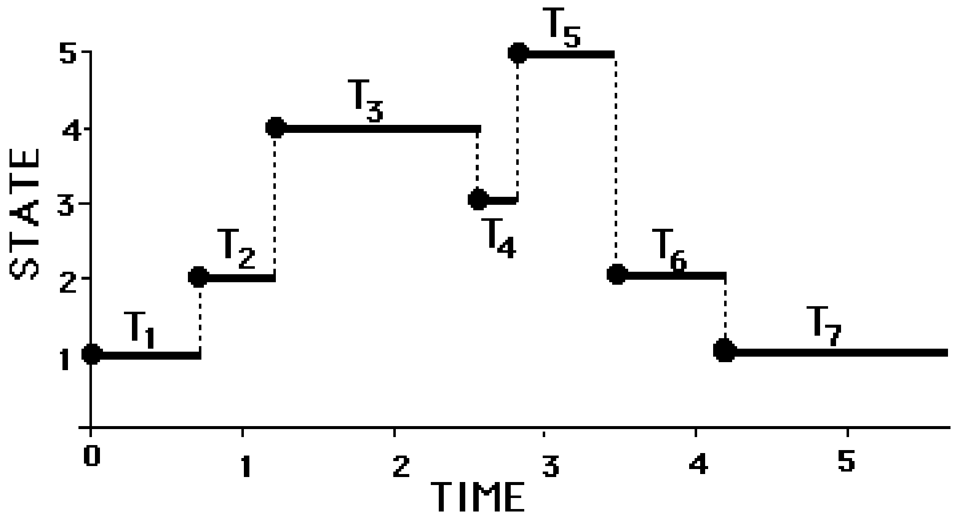
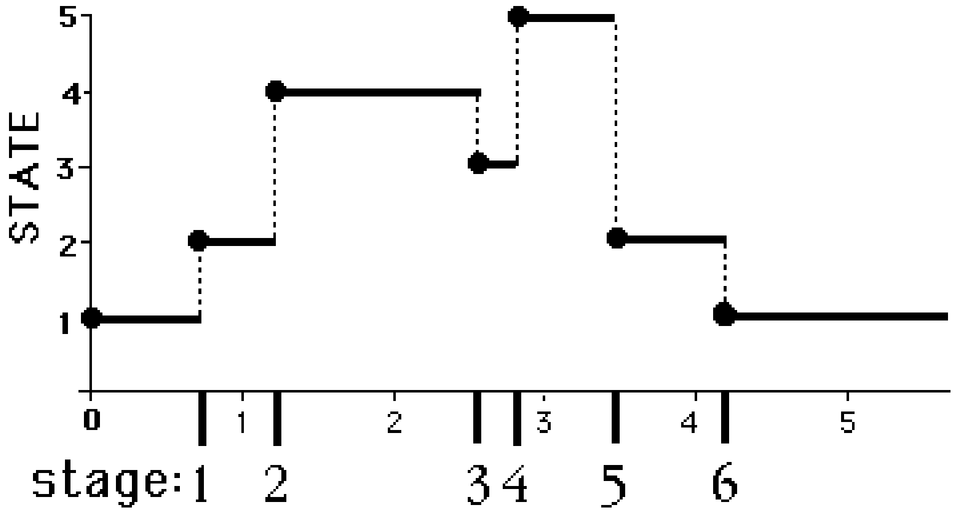
A *continuous-time Markov Chain* (CTMC) may change its state at *any* point in time:



The length of time spent in a state before a transition has the *exponential* distribution:



The ***embedded*** (discrete-time) Markov chain derived from a CTMC:



A **Continuous-time Markov Chain**

is a stochastic process $\{X(t): t \geq 0\}$ where

- $X(t)$ can have values in $\mathbf{S} = \{0, 1, 2, 3, \ldots\}$
- Each time the process enters a state **i**, the amount of time it spends in that state before making a transition to another state has an exponential distribution with mean time $1/\lambda_i$
- When leaving state **i**, the process moves to a state **j** with probability p_{ij} where $p_{ii} = 0$ and $\sum_{j=0}^{M} p_{ij} = 1$
- The next state to be visited after state **i** is independent of the length of time spent in state **i**

Transition Probabilities

$$p_{ij}(t) = P\{ X(t+s) = j \mid X(s) = i \}$$

Continuous at t=0, with

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

$$P(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \cdots \\ p_{21}(t) & \ddots & \\ \vdots & & \end{bmatrix}$$

Transition matrix is a function of time!

Transition Intensity

$$\lambda_j = -\frac{d}{dt} p_{jj}(0)$$ (rate at which the process leaves state j when it is in state j)

$$\lambda_{ij} = \frac{d}{dt} p_{ij}(0) = \lambda_i \, p_{ij}$$ (transition rate into state j when the process is in state i)

The process, starting in state i, spends an amount of time in that state having exponential distribution with rate λ_i. It then moves to state j with probability

$$p_{ij} = \frac{\lambda_{ij}}{\lambda_i} \quad \forall\ i,j$$

$$1 = \sum_{j=1}^{n} p_{ij} = \sum_{j=1}^{n} \frac{\lambda_{ij}}{\lambda_i} = \frac{\sum_{j=1}^{n} \lambda_{ij}}{\lambda_i} \Rightarrow \lambda_i = \sum_{j=1}^{n} \lambda_{ij}$$

Chapman-Kolmogorov Equation

$$p_{ij}(t+s) = \sum_{k\in S} p_{ik}(t)\, p_{kj}(s), \qquad \forall i,j\in S, \ \forall s,t\geq 0$$

Since p (t) is a continuous function,

$$p_{ij}(\Delta t) = p_{ij}(0) + \frac{d}{dt}p_{ij}(0)\Delta t + o(\Delta t^2)$$

But we have defined $\lambda_{ij} = \frac{d}{dt}p_{ij}(0)$

For $i \neq j$: $p_{ij}(\Delta t) = p_{ij}(0) + \lambda_{ij}\Delta t + o(\Delta t^2)$

$\approx \lambda_{ij}\Delta t$ for small Δt

For $i = j$: $p_{ii}(\Delta t) = p_{ii}(0) + \lambda_{ii}\Delta t + o(\Delta t^2)$

$\approx 1 + \lambda_{ii}\Delta t$ for small Δt

From the Chapman-Kolmogorov equation,

$$p_{ij}(t+\Delta t) = \sum_{k} p_{ik}(t)\, p_{kj}(\Delta t)$$

$$= p_{ij}(t)p_{jj}(\Delta t) + \sum_{k\neq j} p_{ik}(t)p_{kj}(\Delta t)$$

$$= p_{ij}(t)\left[1 + \lambda_{jj}\Delta t + o(\Delta t^2)\right] + \sum_{k\neq j} p_{ik}(t)\left[\lambda_{kj}\Delta t + o(\Delta t^2)\right]$$

$$p_{ij}(t+\Delta t) = p_{ij}(t) + \left[\sum_k p_{ik}(t)\,\lambda_{kj}\right]\Delta t + \left[\sum_k p_{ik}(t)\right] o(\Delta t^2)$$

$$\frac{p_{ij}(t+\Delta t) - p_{ij}(t)}{\Delta t} = \sum_k p_{ik}(t)\,\lambda_{kj} + \left[\sum_k p_{ik}(t)\right]\frac{o(\Delta t^2)}{\Delta t}$$

Taking the limit as $\Delta t \longrightarrow 0$

$$\boxed{\frac{d}{dt}\,p_{ij}(t) = \sum_k p_{ik}(t)\,\lambda_{kj} \quad \forall\ \mathbf{i,j}}$$

The process is described by the system of differential equations:

$$\frac{d}{dt} p_{ij}(t) = \sum_{k} p_{ik}(t)\, \lambda_{kj} \quad \forall\ \mathbf{i,j}$$

or

$$\frac{d}{dt} P(t) = P(t)\, \Lambda$$

$$\sum_j p_{ij}(t) = 1 \qquad \forall\ i, t$$

$$\Longrightarrow \quad \frac{d}{dt} \sum_j p_{ij}(t) = \frac{d}{dt}(1) = 0$$

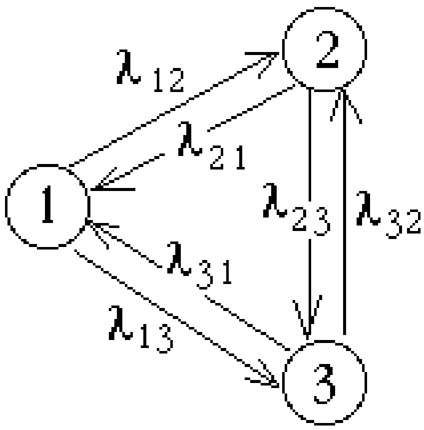
$$\Longrightarrow \quad \sum_j \frac{d}{dt} p_{ij}(t) = 0$$

$$\Longrightarrow \quad \sum_j \lambda_{ij} = 0$$

That is, the sum of each row of Λ is zero!

Example

$$\Lambda = \begin{bmatrix} -(\lambda_{12}+\lambda_{13}) & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & -(\lambda_{21}+\lambda_{23}) & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & -(\lambda_{31}+\lambda_{32}) \end{bmatrix}$$



The sum of each row of Λ must equal zero!

Steadystate Probabilities

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

(independent of the initial state i)

Must be nonnegative and satisfy

$$\sum_{i=1}^{n} \pi_i = 1$$

What other equations are needed to determine π *?*

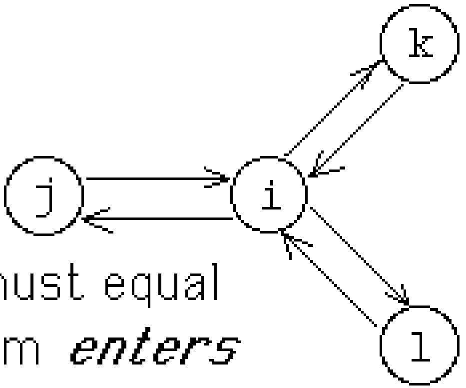
Steadystate Probabilities

In the case of discrete-time Markov chains, we used the equations $\pi = \pi P$

i.e., $\pi_j = \sum_{i=1}^{n} \pi_i \, p_{ij} \;\forall\, j$

In the case of continuous-time Markov chains, we use what are called "Balance" equations.

Balance Equations



For each state i, the rate at which the system *leaves* the state must equal the rate at which the system *enters* the state:

$$\lambda_i \pi_i = \lambda_{ji}\pi_j + \lambda_{ki}\pi_k + \lambda_{li}\pi_l$$

Balance Equations

$$\Big(\sum_{j\neq i} \lambda_{ij}\Big)\pi_i = \sum_{k\neq i} \lambda_{ki}\,\pi_k \qquad \forall\, i$$

transition rates **from** state i

transition rates **into** state i

steadystate distribution is computed by solving this system of equations

$$\begin{cases} \Big(\sum_{j\neq i} \lambda_{ij}\Big)\pi_i = \sum_{k\neq i} \lambda_{ki}\,\pi_k & \forall\, i \\ \sum_{i=1}^{n} \pi_i = 1 & \end{cases}$$

An alternate derivation of the steady-state conditions begins with the differential equation describing the process:

$$\boxed{\frac{d}{dt}\, p_{ij}(t) = \sum_k p_{ik}(t)\,\lambda_{kj} \quad \forall\ \mathbf{i,j}}$$

Suppose that we take the limit of each side, as $t \rightarrow \infty$

$$\boxed{\frac{d}{dt} p_{ij}(t) = \sum_k p_{ik}(t)\, \lambda_{kj} \quad \forall\ \mathbf{i,j}}$$

$$\Longrightarrow \quad \lim_{t\to\infty} \frac{d}{dt} p_{ij}(t) = \lim_{t\to\infty} \sum_k p_{ik}(t)\, \lambda_{kj}$$

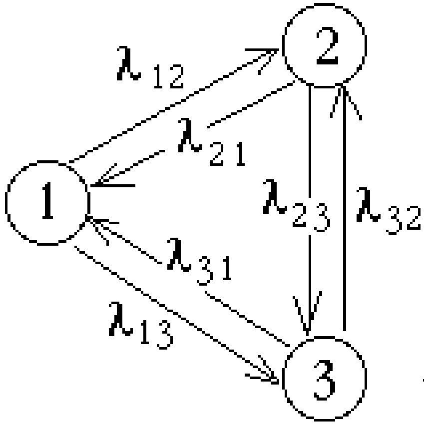
$$\Longrightarrow \quad \frac{d}{dt}\left(\lim_{t\to\infty} p_{ij}(t)\right) = \sum_k \left(\lim_{t\to\infty} p_{ik}(t)\right) \lambda_{kj}$$

$$\Longrightarrow \quad 0 = \sum_k \pi_k\, \lambda_{kj}$$

i.e., $\boxed{\pi\, \Lambda = 0}$

Example

$$\Lambda = \begin{bmatrix} -(\lambda_{12}+\lambda_{13}) & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & -(\lambda_{21}+\lambda_{23}) & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & -(\lambda_{31}+\lambda_{32}) \end{bmatrix}$$

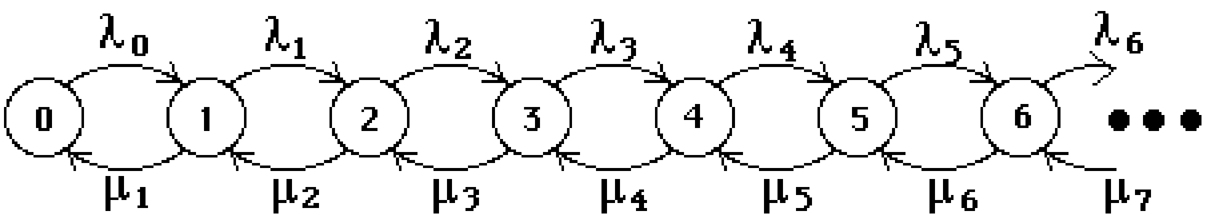


$$\boxed{\pi \Lambda = 0}$$

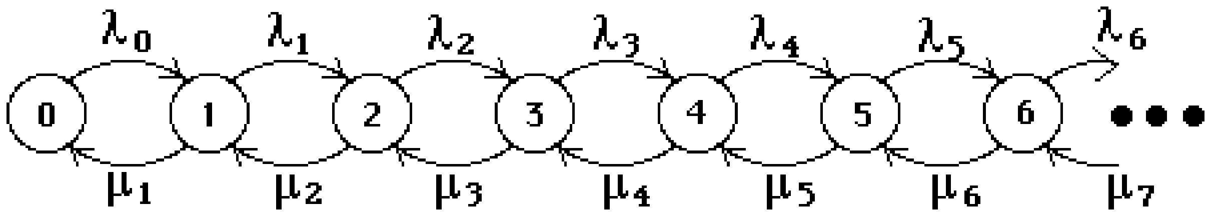
$$\begin{cases} -(\lambda_{12}+\lambda_{13})\pi_1 + \lambda_{21}\pi_2 + \lambda_{31}\pi_3 = 0 \\ \lambda_{12}\pi_1 - (\lambda_{21}+\lambda_{23})\pi_2 + \lambda_{32}\pi_3 = 0 \\ \lambda_{13}\pi_1 + \lambda_{23}\pi_2 - (\lambda_{31}+\lambda_{32})\pi_3 = 0 \end{cases}$$

Birth-Death Process

A birth-death process is a continuous-time Markov chain which models the size of a population; the population increases by 1 ("birth") or decreases by 1 ("death").



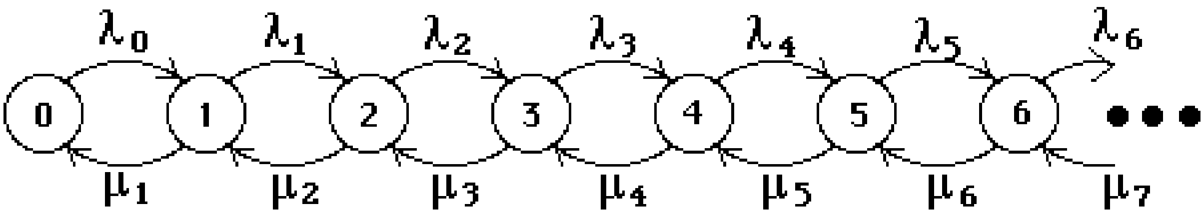
Steady-State Distribution of a Birth-Death Process



Balance Equations:

State 0: $\lambda_0 \pi_0 = \mu_1 \pi_1 \quad \Rightarrow \quad \boxed{\pi_1 = \frac{\lambda_0}{\mu_1} \pi_0}$

Steady-State Distribution of a Birth-Death Process



Balance Equations:

State 1: $(\lambda_1 + \mu_1)\,\pi_1 = \lambda_0\pi_0 + \mu_2\pi_2 \Rightarrow$

$$\pi_2 = \frac{(\lambda_1 + \mu_1)\,\pi_1 - \lambda_0\pi_0}{\mu_2} = \frac{(\lambda_1 + \mu_1)\frac{\lambda_0}{\mu_1}\pi_0 - \lambda_0\pi_0}{\mu_2}$$

$$\Rightarrow \boxed{\pi_2 = \frac{\lambda_1\lambda_0}{\mu_2\mu_1}\pi_0}$$

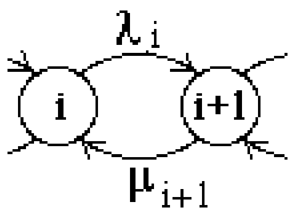
In general,

$$(\lambda_{i-1} + \mu_{i-1})\,\pi_{i-1} = \lambda_{i-2}\pi_{i-2} + \mu_i\pi_i$$

$$\Rightarrow \quad \boxed{\pi_i = \frac{\lambda_{i-1}\cdots\lambda_1\lambda_0}{\mu_i\cdots\mu_2\mu_1}\,\pi_0} \quad i=1,2,3,\ \ldots$$

$$\pi_i = \left(\frac{\lambda_{i-1}}{\mu_i}\right) \cdots \left(\frac{\lambda_1}{\mu_2}\right)\left(\frac{\lambda_0}{\mu_1}\right)\pi_0$$

$$= \rho_{i-1} \cdots \rho_1 \; \rho_0 \, \pi_0 \quad \textbf{where } \rho_i = \frac{\lambda_i}{\mu_{i+1}}$$



ratio of transition rates between adjacent states

Substituting these expressions for π_i into

$$\sum_{i=0}^{\infty} \pi_i = 1$$ yields:

$$\pi_0 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1} \cdots \lambda_1 \lambda_0}{\mu_i \cdots \mu_2 \mu_1} \pi_0 = 1$$

$$\Rightarrow \pi_0 \left[1 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1} \cdots \lambda_1 \lambda_0}{\mu_i \cdots \mu_2 \mu_1} \right] = 1$$

$$\Rightarrow \boxed{\frac{1}{\pi_0} = \left(1 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1} \cdots \lambda_1 \lambda_0}{\mu_i \cdots \mu_2 \mu_1} \right)}$$

Once π_0 is evaluated by computing the reciprocal of this infinite sum, π_i is easily computed for each $i=1, 2, 3, \ldots$

$$\boxed{\frac{1}{\pi_0} = \left(1 + \sum_{i=1}^{\infty} \frac{\lambda_{i-1} \cdots \lambda_1 \lambda_0}{\mu_i \cdots \mu_2 \mu_1}\right)}$$

$$\boxed{\pi_i = \frac{\lambda_{i-1} \cdots \lambda_1 \lambda_0}{\mu_i \cdots \mu_2 \mu_1}\, \pi_0} \quad i=1,2,3, \ldots$$

Examples

- Backup Computer System
- Multiple Failure Modes
- The "Peter Principle"
- Gasoline Station
- Ticket Sales by Phone

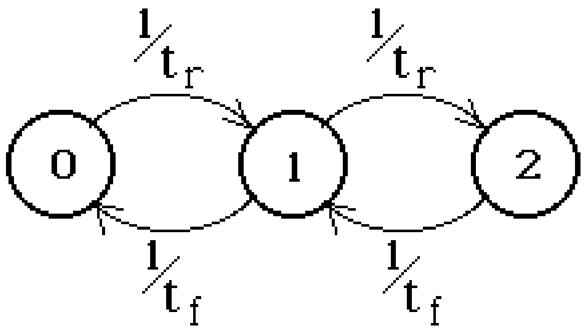
Example

An airlines reservation system has 2 computers, one on-line and one standby. The operating computer fails after an exponentially-distributed duration having mean t_f and is then replaced by the standby computer.

There is one repair facility, and repair times are exponentially-distributed with mean t_r.

What fraction of the time will the system fail, i.e., both computers having failed?

Let $X(t)$ = number of computers in operating condition at time t. Then $X(t)$ is a birth-death process.



Note that the birth rate in state 2 is zero!

$$\frac{1}{\pi_0} = 1 + \frac{1/t_r}{1/t_f} + \left(\frac{1/t_r}{1/t_f}\right)^2$$

$$\frac{1}{\pi_0} = 1 + \frac{t_f}{t_r} + \left(\frac{t_f}{t_r}\right)^2$$

$$\pi_0 = \frac{t_r^2}{t_r^2 + t_r t_f + t_f^2}$$

probability that both computers have failed

Suppose that $\frac{t_f}{t_r} = 10$, i.e., the average repair time is 10% of the average time between failures:

$$\frac{1}{\pi_0} = 1 + 10 + 100 = 111$$

$$\pi_0 = \frac{1}{111} = 0.009009$$

Then both computers will be simultaneously out of service 0.9% of the time.

Example: Multiple Failure Modes

A production system consists of 2 machines, both of which may operate simultaneously, and a single repair facility.

The machines each fail randomly, with time between failures having exponential distribution and mean T hours.

Repair times are also exponentially distributed, but the mean repair time depends upon whether the failure was "regular" or "severe".

The fraction of regular failures is p, and the corresponding mean repair time is t_r. The fraction of severe failures is $q=1-p$, and the mean repair time is t_s.

Let $T=10$ hours, $p=90\%$, $t_r=1$ hour, $t_s=5$ hours.

What is the average number of machines in operation?

Markov Chain Model

States

(0,0): both machines operational

(r,0): regular repair in progress, none waiting

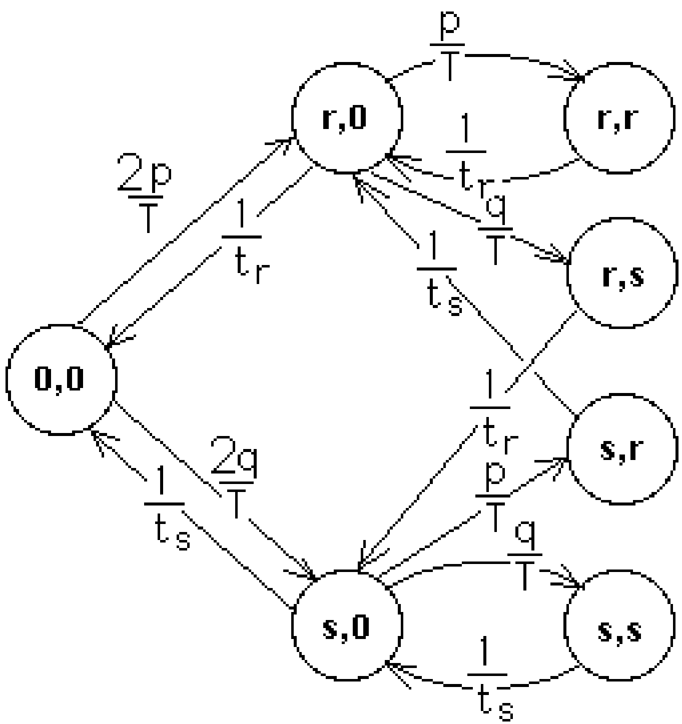
(r,r): regular repair in progress, regular waiting

(r,s): regular repair in progress, severe waiting

(s,0): severe repair in progress, none waiting

(s,r): severe repair in progress, regular waiting

(s,s): severe repair in progress, severe waiting

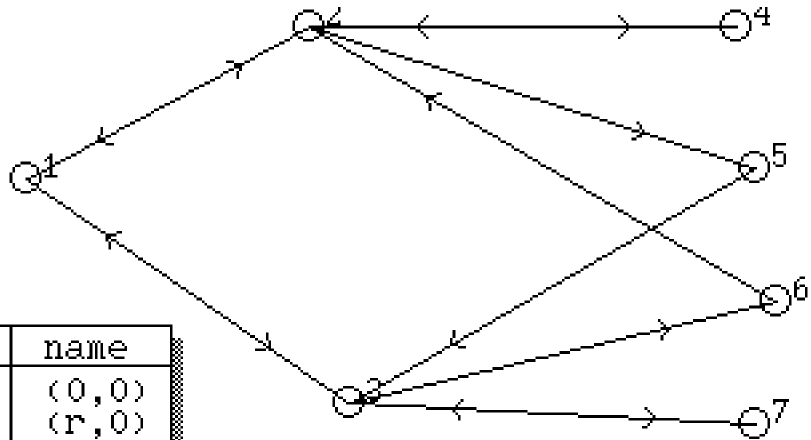


$\frac{p}{T}$	**rate of regular failures**
$\frac{q}{T}$	**rate of severe failures**
$\frac{1}{t_r}$	**regular repair rate**
$\frac{1}{t_s}$	**severe repair rate**

Transition rate matrix

from \ to	1	2	3	4	5	6	7
1	¯0.2	0.18	0.02	0	0	0	0
2	1	¯1.1	0	0.09	0.01	0	0
3	0.2	0	¯0.3	0	0	0.09	0.01
4	0	1	0	¯1	0	0	0
5	0	0	1	0	¯1	0	0
6	0	0.2	0	0	0	¯0.2	0
7	0	0	0.2	0	0	0	¯0.2

i	name
1	(0,0)
2	(r,0)
3	(s,0)
4	(r,r)
5	(r,s)
6	(s,r)
7	(s,s)



i	name
1	(0,0)
2	(r,0)
3	(s,0)
4	(r,r)
5	(r,s)
6	(s,r)
7	(s,s)

Steadystate Distribution

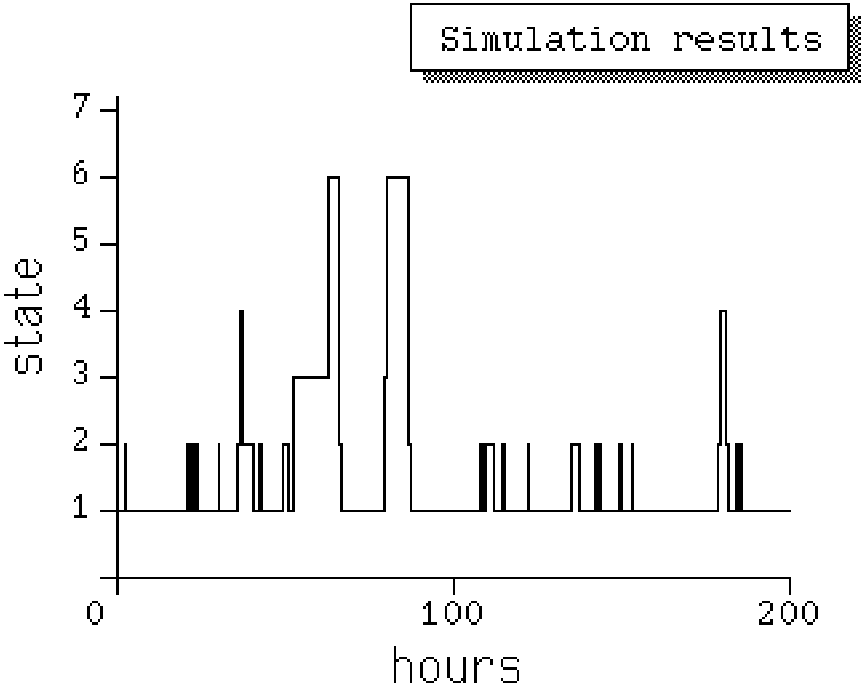
i	state	Pi
1	(0,0)_____	0.7596253902
2	(r,0)_____	0.1404786681
3	(s,0)_____	0.05723204995
4	(r,r)_____	0.01264308012
5	(r,s)_____	0.001404786681
6	(s,r)_____	0.02575442248
7	(s,s)_____	0.002861602497

i	state	Pi	C	Pi×C
1	(0,0)_____	0.7596253902	2	1.51925078
2	(r,0)_____	0.1404786681	1	0.1404786681
3	(s,0)_____	0.05723204995	1	0.05723204995
4	(r,r)_____	0.01264308012	0	0
5	(r,s)_____	0.001404786681	0	0
6	(s,r)_____	0.02575442248	0	0
7	(s,s)_____	0.002861602497	0	0

The average cost/period in steady state is 1.716961498

$1.716961498 \div 2 = 0.858480749$

In steady state, the system will operate at approximately 85.8% of capacity.


Simulation results
state
7
6
5
4
3
2
1
0
100
200
hours

Simulation results

Random seed: 675247
Initial state: 1

state	1	2	3	4	5	6	7
# visits	23	24	2	2	0	2	0
time in state	19.814	138.53	15.713	1.661	0	10.381	0
% total time	9.907	69.264	7.856	0.830	0	5.190	0

Example: The Peter Principle

The draftsman position at a large engineering firm can be occupied by a worker at any of three levels:

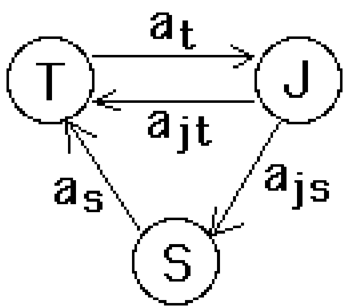
T= Trainee
J = Junior Draftsman
S = Senior Draftsman

Assume that a Trainee stays at a rank for an exponentially-distributed length of time (with parameter a_t) before being promoted to Junior Draftsman.

A Junior Draftsman stays at that level for an exponentially-distributed length of time (with parameter $a_j = a_{jt} + a_{js}$). Then he either leaves the position and is replaced by a Trainee (with probability a_{jt}/a_j), or is promoted to a Senior Draftsman (with probability a_{js}/a_j).

Senior Draftsmen remain in that position an exponentially-distributed length of time (with parameter a_S) before resigning or retiring, in which case they are replaced by a Trainee.

The rank of a person in a draftsman's position may be modeled as a continuous-time Markov chain

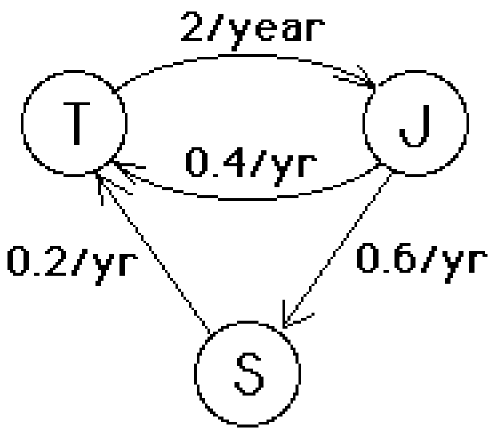


For example, suppose that the mean time in the three ranks are:

State	Mean Time
T	0.5 years
J	1 year
S	5 years

and that a Junior Draftsman leaves and is replaced by a Trainee with probability 40% and is promoted with probability 60%

Steadystate Distribution



balance eqns.
$$\begin{cases} 2\,\pi_t = 0.4\,\pi_j + 0.2\,\pi_s \\ 1\,\pi_j = 2\,\pi_t \\ 0.2\,\pi_s = 0.6\,\pi_j \end{cases}$$

$$\pi_t + \pi_j + \pi_s = 1$$

$$\Rightarrow \begin{cases} \pi_t = 0.11 \\ \pi_j = 0.22 \\ \pi_s = 0.67 \end{cases}$$

i.e., 11% of draftsmen are trainees, 22% are at junior rank, & 67% at senior rank

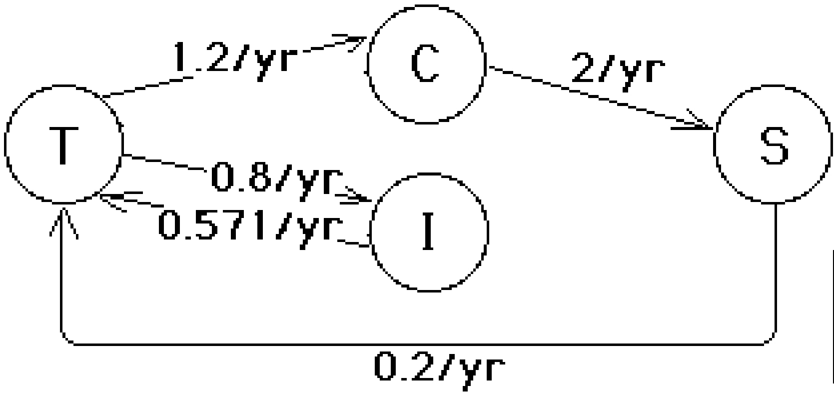
The duration that people spend in any given rank is not exponentially distributed in general. A **bimodal** distribution is often observed in which many people leave (are promoted) rather quickly, while others persist for a substantial time.

The "*Peter Principle*" asserts that a worker is promoted until first reaching a position in which he or she is incompetent. When this happens, the worker stays in that job until retirement.

Let's modify the above model by classifying 60% of the Junior Draftsmen as **C**ompetent and 40% as **I**ncompetent, represented by states **C** and **I**, respectively.

Suppose that incompetent draftsmen stay at that rank until quitting or retirement (an average of 1.75 years) and competent draftsment are promoted (after an average of 0.5 years), so that the average time spent in the rank is still

$$(0.6)(.5) + (0.4)(1.75) = 1 \text{ year}$$



steadystate distribution

balance eqns.
$$\begin{cases} 1.2\,\pi_t = 0.571\,\pi_i + 0.2\,\pi_s \\ 2\,\pi_c = 1.2\,\pi_t \\ 0.571\,\pi_i = 0.8\,\pi_t \\ 0.2\,\pi_s = 2\,\pi_c \end{cases}$$

$$\pi_t + \pi_c + \pi_i + \pi_s = 1$$

$$\Longrightarrow \begin{cases} \pi_t = 0.111 \\ \pi_c = 0.067 \\ \pi_i = 0.155 \\ \pi_s = 0.667 \end{cases}$$

$$\left\{\begin{array}{l}\pi_t = 0.111 \\ \left.\begin{array}{l}\pi_c = 0.067 \\ \pi_i = 0.155\end{array}\right\} \text{total} = 0.222 \ \text{ as before} \\ \pi_s = 0.667\end{array}\right.$$

$$\left(\frac{0.067}{0.222} = 30\%\right)$$

While only 40% of the draftsmen promoted to junior rank are incompetent, we see that the rank of junior draftsmen is 70% filled with incompetent persons!

A gasoline station has only one pump.

Cars arrive at the rate of 20/hour.

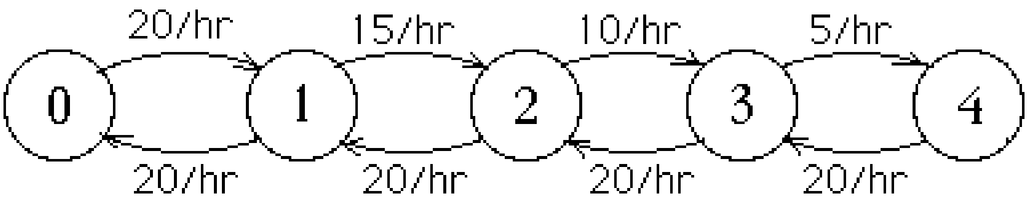
However, if the pump is already in use, these potential customers may "balk", i.e., drive on to another gasoline station.

> If there are n cars already at the station, the probability that an arriving car will balk is $n/4$, for $n=1,2,3,4$, and 1 for $n>4$.

Time required to service a car is exponentially distributed, with mean = 3 minutes.

What is the expected waiting time of customers?

"Birth/death" model:



$$\frac{1}{\pi_0} = 1 + \frac{20}{20} + \frac{20}{20}\times\frac{15}{20} + \frac{20}{20}\times\frac{15}{20}\times\frac{10}{20} + \frac{20}{20}\times\frac{15}{20}\times\frac{10}{20}\times\frac{5}{20}$$

$$= 1 + 1 + 0.75 + 0.375 + 0.09375 = 3.21875$$

$$\pi_0 = 0.3106796$$

Steady State Distribution

$$\pi_0 = 0.3106796,$$

$$\pi_1 = \pi_0 = 0.3106796,$$

$$\pi_2 = 0.75\pi_0 = 0.2330097,$$

$$\pi_3 = 0.375\pi_0 = 0.1165048,$$

$$\pi_4 = 0.09375\pi_0 = 0.0291262$$

Average Number in System

$$L = \sum_{i=0}^{4} i\,\pi_i$$

$$= 0.3106796 + 2(0.2330097)$$

$$+ 3(0.1165048) + 4(0.0291262)$$

$$= 1.2427183$$

Average Arrival Rate

$$\bar{\lambda} = \sum_{i=0}^{4} \lambda_i \pi_i$$

$$= (0.3106796)\times 20/\text{hr} + (0.3106796)\times 15/\text{hr}$$
$$+ (0.2330097)\times 10/\text{hr} + (0.1165048)\times 5/\text{hr}$$
$$+ (0.0291262)\times 0/\text{hr}$$
$$= 13.786407/\text{hr}$$

Average Time in System

$$W = L/\lambda = \frac{1.2427183}{13.786407/hr}$$

$$= 0.0901408 \text{ hr.} = 5.40844504 \text{ minutes}$$

Hancher Auditorium has 2 ticket sellers who answer phone calls & take incoming ticket reservations, using a single phone number.

In addition, 2 callers can be put "on hold" until one of the two ticket sellers is available to take the call.

If all 4 phone lines are busy, a caller will get a busy signal, and waits until later before trying again.

Calls arrive at an average rate of 2/minute, and ticket reservations service time averages 20 sec. and is exponentially distributed.

What is...

- the fraction of the time that each ticket seller is idle?
- the fraction of customers who get a busy signal?
- the average waiting time ("on hold")?