DUBLIN CITY UNIVERSITY

CA4011 Operations Research

B. Markov Chains, Queues and related topics

Contents

1. Markov Chains, Processes & related topics

1.1 Some definitions

A stochastic process, applied to a system with distinct states, means a process in which probabilities are associated with entering a particular state. The probabilities usually depend on the previous history of the system. If the set of distinct states is discrete (rather than continuous) then the term STOCHASTIC CHAIN is often used.

A Markov Process is a particular kind of stochastic process where the probability of entering a particular state depends only on the **last** state occupied as well as on the matrix governing the process. The term MARKOV CHAIN is used when the set of states is discrete.

It may be useful to refer to the chapter on Markov chains in (Taha, 2007). In addition there are some references given throughout section 1. There is an interesting application of Markov Chains in (McGarigal, n.d.) which may be helpful as a visual introduction to the topic.

1.2 Example 1 (data)

Newspaper readership one year ago:-

Paper A 200, Paper B 400, Paper C 100

Readership this year

		A	B	C	Totals (one year ago)
Last	A	160	20	20	200
Year	B	40	280	80	400
	C	10	30	60	100
Totals (this year)		210	330	160	

For example, 160 of 200 reading A 1 year ago are reading A now, 20 switched to B and 20 switched to C.

Let $\mathbf{p}(0) = [200\ 400\ 100]$ and $\mathbf{p}(1) = [210\ 330\ 160]$.

1.3 Transition matrix; Markov assumption

P_{ij} = probability, given now in state i, of being in state j one period later. These probabilities form a matrix **P** called the **transition matrix**

For Example 1, we obtain **P** by dividing by the row totals. For example, $P_{AA} = 0.8$ = probability, given reading A a year ago, of reading A now.

$$\underline{P} = \begin{array}{c|ccc} & A & B & C \\ \hline A & 0.8 & 0.1 & 0.1 \\ B & 0.1 & 0.7 & 0.2 \\ C & 0.1 & 0.3 & 0.6 \end{array}$$

Markov assumption: $P(X_t = j \mid X_{t-1} = i, X_{t-2} = k, \ldots\ldots) = P(X_t = j \mid X_{t-1} = i)$

that is, the. system does not 'remember' back more than one time period.

1.4 Use of transition matrices

For example 1, we had the initial state and the state after one time period: **p**(0) = [200 400 100] and **p**(1) = [210 330 160].

Consider the product

$$[200\ 400\ 100] \begin{vmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.3 & 0.6 \end{vmatrix}$$

= [210 330 160] = **p**(1)

In fact, we have in general that **p**(0)**P** = **p**(1)

1.5 Stationarity assumption

If we further assume **stationarity** (i.e. that transition probabilities remain constant)

then $\mathbf{p}(2) = \mathbf{p}(1)\mathbf{P}$

In general $\mathbf{p}(t) = \mathbf{p}(t-1)\mathbf{P}$

Therefore, $\mathbf{p}(t) = \mathbf{p}(0)\mathbf{P}^t$

[e.g. if t =4, then $\mathbf{p}(4) = \mathbf{p}(3)\mathbf{P} = \mathbf{p}(2)\mathbf{P}^2 = \mathbf{p}(1)\mathbf{P}^3 = \mathbf{p}(0)\mathbf{P}^4$]

Note: A Markov chain with stationarity is said to be homogeneous.

Applying this to Example 1, we find the readership after three years: **p**(3) = **p**(0)**P**^{3} =

$$[200\ 400\ 100]\begin{vmatrix}0.8 & 0.1 & 0.1\\0.1 & 0.7 & 0.2\\0.1 & 0.3 & 0.6\end{vmatrix}\begin{vmatrix}0.8 & 0.1 & 0.1\\0.1 & 0.7 & 0.2\\0.1 & 0.3 & 0.6\end{vmatrix}\begin{Vmatrix}0.8 & 0.1 & 0.1\\0.1 & 0.7 & 0.2\\0.1 & 0.3 & 0.6\end{Vmatrix}$$

=

$$[200\ 400\ 100]\begin{vmatrix}0.8 & 0.1 & 0.1\\0.1 & 0.7 & 0.2\\0.1 & 0.3 & 0.6\end{vmatrix}\begin{vmatrix}0.66 & 0.18 & 0.16\\0.17 & 0.56 & 0.27\\0.17 & 0.40 & 0.43\end{vmatrix}$$

=

$$[200\ 400\ 100]\begin{vmatrix}0.562 & 0.240 & 0.198\\0.219 & 0.490 & 0.291\\0.219 & 0.426 & 0.355\end{vmatrix}$$

=

[221.9 286.6 191.5]

1.6 Steady state, related conditions & definitions

1.6.1 How to determine the steady-state

For Example 1, we could proceed similarly to find

p(10) = **p**(0)**P**^{10} = [232.4 272.9 194.7] [*10 matrix mults!*]

p(15) =**p**(0)**P**^{15} = [233.2 272.3 194.5] [*15 matrix mults!*]

p(20) =**p**$(0)\underline{P}^{20}$ = [233.2 272.3 194.5] [*20 matrix mults!*]

Of course, this is very laborious and would **never be done** normally. We do it here just to show that the system settles down, that is reaches a **steady state**. (e.g. **p**(15)= **p**(20) to the precision shown). Actually, we denote the **proportions** by $\boldsymbol{\pi}$, that is

	A	B	C
Readers	233.2	272.3	194.5
Proportions	233.2/700 = 0.333 (π_1)	0.389 (π_1)	0.278 (π_1)

In fact, if there is a steady state we can calculate it directly without carrying out large numbers of matrix multiplications, as follows.

As stated, we denote the steady state proportions by $\boldsymbol{\pi} = [\pi_1\ \pi_2\ \pi_3\ \ldots\ \pi_n]$

Steady state condition in general:

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} \qquad (\text{or } \boldsymbol{\pi}^T = \mathbf{P}^T\boldsymbol{\pi}^T)$$

i.e. the proportions remain the same if we post-multiply by the transition matrix.

Note: Superscript "^{T}" denotes "transpose".

For Example 1, this becomes

$$[\pi_1\ \pi_2\ \pi_3] = [\pi_1\ \pi_2\ \pi_3] \begin{vmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.3 & 0.6 \end{vmatrix}$$

=>

$$\rightarrow \quad \pi_1 = 0.8\pi_1 + 0.1\pi_2 + 0.1\pi_3$$
$$\pi_2 = 0.1\pi_1 + 0.7\pi_2 + 0.3\pi_3$$
$$\rightarrow \quad \pi_3 = 0.1\pi_1 + 0.2\pi_2 + 0.6\pi_3$$

We also have the condition that the sum of proportions is one, i.e.

$$\pi_1 + \pi_2 + \pi_3 = 1$$

This is a system of 4 equations and 3 unknowns. First take the first 3 (homogeneous) equations and re-write them as,

$$\begin{array}{l} -2\pi_1 + \pi_2 + \pi_3 = 0 \\ \pi_1 - 3\pi_2 + 3\pi_3 = 0 \\ \pi_1 + 2\pi_2 - 4\pi_3 = 0 \end{array} \quad => \quad \begin{array}{r} -5\pi_2 + 7\pi_3 = 0 \\ \pi_1 - 3\pi_2 + 3\pi_3 = 0 \\ 5\pi_2 - 7\pi_3 = 0 \end{array}$$

We can see that two of the equations are the same (i.e. one can be eliminated). Then, it is easy to see that $\pi_2 = (7/5)\pi_3$ and then that $\pi_1 = (6/5)\pi_3$.

Finally, applying the additional condition we have

$$(6/5)\pi_3 + (7/5)\pi_3 + \pi_3 = 1 => \pi_3 = 5/18 = 0.278$$

Hence, the solution is $\pi_1 = 0.333 \quad \pi_2 = 0.389 \quad \pi_3 = 0.278$

We got this before by matrix multiplications but this method via equations is exact.

1.6.2 Some related terminology and definitions

Note: **The rows of transition matrix P add to 1 always**. However, if, in addition, its columns also add to 1, then **P** is said to be <u>Doubly Stochastic.</u> In this special case it can be shown that the steady state probabilities are just $\pi_1=\pi_2=\ldots=\pi_n=1/n$.

<u>Definition of Irreducible Transition Matrix</u>: If there exists some n for which $\mathbf{P}^n$ has every element > 0 (i.e. every state can be reached from every other state at some stage), the transition matrix is said to be <u>Irreducible</u>.

<u>Steady-State Condition</u>: It is possible to calculate steady state conditions if the transition matrix is irreducible.

Absorbing states: State i is absorbing (trapping) if P_{ii}=1 (and P_{ij}=0 for all $i \neq j$) → impossible to leave state i.

Absorbing Markov chain: A Markov chain is absorbing if it has one or more absorbing states. If it has one absorbing state (state i, say) then the steady state conditions are $\pi_i = 1.0$ and $\pi_j = 0$ for all $i \neq j$.

No steady-state solution– Multiple absorbing states: If there is more than one absorbing state, we cannot calculate steady state conditions.

1.6.3 Algebraic Eigenvalue Problem applied to Markov chains

You may recall (from linear algebra)that this problem is to find non-zero scalars (real numbers) λ and corresponding vectors **x** to satisfy the equation:

$\mathbf{A}_{nxn}\mathbf{x}_{nx1} = \lambda\mathbf{x}_{nx1}$ or $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ (where **I** is the nxn unit matrix)

Scalars λ are called eigenvalues and vectors **x** eigenvectors. The full set of eigenvalues is the **spectrum** of matrix **A**, and the largest eigenvalue in absolute value is called the **spectral radius** of **A**.

If all eigenvalues are different then there are n of them.

Eigenvectors are determined to within a scalar multiple, i.e. if **x** is an eigenvector corresponding to eigenvalue λ then so is k**x** where k is a non-zero scalar.

We had the steady state condition:

$$\boldsymbol{\pi} = \boldsymbol{\pi}\,\mathbf{P} \; (or\; \boldsymbol{\pi}^T = \mathbf{P}^T\boldsymbol{\pi}^T)$$

We can see that $\boldsymbol{\pi}^T = \mathbf{P}^T\boldsymbol{\pi}^T$ is an eigenvalue problem with eigenvalue equal to 1.

In fact, it can be shown that **P** and $\mathbf{P}^T$ have the same eigenvalues, from which it follows that **P** also has an eigenvalue 1.

1.7 Example 2

Suppose that the 2003 state of land use in a city of 50 square miles of non-vacant land was

I	Residential	30%
II	Commercial	20%
III	Industrial	50%

Find the states in 2008 and 2013, assuming that the transition probabilities for 5-year intervals are given by the matrix P,

	To I	To II	To III
From I	0.8	0.1	0.1
From II	0.1	0.7	0.2
From II	0	0.1	0.9

Solution:

2008 I (res) 0.3 *0.8 + 0.2*0.1 + 0.5*0.0 = 0.26

II(comm) 0.3 *0.1 + 0.2*0.7 + 0.5*0.1 = 0.22

III(ind) 0.3 *0.1 + 0.2*0.2 + 0.5*0.9 = 0.52

or, in matrix form, **p**(2008)=**p**(2003)**P** = [0.26 0.22 0.52]

Similarly, **p**(2013)=**p**(2008)**P** =

$$\begin{vmatrix} 0.26 & 0.22 & 0.52 \end{vmatrix} \begin{vmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0 & 0.1 & 0.9 \end{vmatrix}$$

$$= [0.23 \;\; 0.232 \;\; 0.538]$$

$$= \underline{p}(2003)\underline{P}^2. \qquad (\textit{STATIONARITY ASSUMED})$$

Note: We can calculate powers of the "$\underline{P}$" matrix. For example

$$\underline{P}^2 = \begin{vmatrix} 0.65 & 0.16 & 0.19 \\ 0.15 & 0.52 & 0.33 \\ 0.01 & 0.16 & 0.83 \end{vmatrix}$$

$$\underline{P}^4 = \begin{vmatrix} 0.4484 & 0.2176 & 0.334 \\ 0.1788 & 0.3472 & 0.474 \\ 0.0388 & 0.2176 & 0.7436 \end{vmatrix}$$

$$\underline{P}^8 = \begin{vmatrix} 0.2529 & 0.2458 & 0.5013 \\ 0.1606 & 0.2630 & 0.5768 \\ 0.0852 & 0.2458 & 0.6690 \end{vmatrix}$$

$$\underline{P}^{16} = \begin{vmatrix} 0.1461 & 0.2499 & 0.6039 \\ 0.1319 & 0.2502 & 0.6179 \\ 0.1180 & 0.2499 & 0.6321 \end{vmatrix}$$

$$\underline{P}^{64} = \begin{vmatrix} 0.125 & 0.25 & 0.625 \\ 0.125 & 0.25 & 0.625 \\ 0.125 & 0.25 & 0.625 \end{vmatrix}$$

So, in the steady-state, the proportions are 1/8 residential, ¼ commercial, and 5/8 industrial.

However, as in previous example, **there is a better, exact way** than having to do many matrix multiplications. We denote the steady state proportions by $\boldsymbol{\pi} = [\pi_1\ \pi_2\ \pi_3\ \pi_n]$

Steady state conditions in general: $\pi = \pi\ \boldsymbol{P}$

For the example, this becomes

$$[\pi_1\ \pi_2\ \pi_3] = [\pi_1\ \pi_2\ \pi_3] \begin{vmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0 & 0.1 & 0.9 \end{vmatrix}$$

Check (***exercise***) that this implies that $\pi_2 = (2/5)\ \pi_3$, $\pi_1 = (1/5)\ \pi_3$. In addition, $\pi_1 + \pi_2 + \pi_3 = 1 \Rightarrow \pi_1 = 1/8$, $\pi_2 = 2/8$ and $\pi_3 = 5/8$, i.e. proportions are 0.125, 0.250 and 0.625 (as found before).

1.8 Example 3: An accounts problem with 2 absorbing states

State	Definition
a	A/c paid
1	1 month overdue
2	2 months overdue
3	3 months overdue
b	Bad debt

The following transition matrix is given, **P** =

Prev\ Next	a	1	2	3	b
a	1	0	0	0	0
1	0.6	0	0.4	0	0
2	0.75	0	0	0.25	0
3	0.5	0	0	0	0.5
b	0	0	0	0	1

We can see that states a and b are absorbing states, which makes sense. We will not have a (single) steady-state solution as **P** is not irreducible (see earlier) and because there is more than one absorbing state.

Suppose that initially the invoices outstanding for 1, 2 and 3 months respectively are €1000, €800 and €200. Then, we can calculate

$\mathbf{S}_0$ = [0 1000 800 200 0]

$\mathbf{S}_1 = \mathbf{S}_0\mathbf{P}$ = [1300 0 400 200 100]

$\mathbf{S}_2 = \mathbf{S}_1\mathbf{P} = [1700 \quad 0 \quad 0 \quad 100 \quad 200]$

$\mathbf{S}_3 = \mathbf{S}_2\underline{P} = [1750 \quad 0 \quad 0 \quad 0 \quad 250]$

Interpretation:

- Of the €2000 outstanding, after 3 months €1750 is paid and €250 is written off as a bad debt.
- As "a" and "b" are absorbing states there cannot be further change (which makes sense!).

Note: We next introduce the idea of a "**fundamental matrix**", a concept basic to handling problems with absorbing states. We use the above data (Example 3) to illustrate.

1.9 Fundamental matrix

First we permute the transition matrix so that the absorbing state rows and columns are adjacent:

Prev\ Next	a	b	1	2	3
a	1	0	0	0	0
b	0	1	0	0	0
1	0.6	0	0	0.4	0
2	0.75	0	0	0	0.25
3	0.5	0.5	0	0	0

We partition the matrix as shown (by the dotted lines) to identify 4 sub-matrices:

I = Identity matrix (in this example 2x2 as there are 2 absorbing states)

O = Matrix of zeros (2x3 as there is zero probability of leaving an absorbing state)

R = Probabilities of being absorbed in next period (3x2 for 3 **non-absorbing** states)

Q = Transition probabilities between non-absorbing states (3x3 in this case)

In general, **P** has the form

$$\underline{P} = \left[\begin{array}{c:c} \underline{I} & \underline{O} \\ \hdashline \underline{R} & \underline{Q} \end{array}\right]$$

We use this form to multiply **P** by itself repeatedly to get the **fundamental matrix F**:

$$\underline{P}^2 = \left|\begin{array}{c:c} \underline{I} & \underline{O} \\ \hdashline \underline{R} + \underline{Q}\underline{R} & \underline{Q}^2 \end{array}\right|$$

$$\underline{P}^3 = \left|\begin{array}{c:c} \underline{I} & \underline{O} \\ \hdashline (\underline{I} + \underline{Q} + \underline{Q}^2)\underline{R} & \underline{Q}^3 \end{array}\right|$$

$$\underline{P}^4 = \left|\begin{array}{c:c} \underline{I} & \underline{O} \\ \hdashline (\underline{I} + \underline{Q} + \underline{Q}^2 + \underline{Q}^3)\underline{R} & \underline{Q}^4 \end{array}\right|$$

...

$$\underline{P}^t = \left|\begin{array}{c:c} \underline{I} & \underline{O} \\ \hdashline \sum_{i-0}^{t-1} \underline{Q}^i\underline{R} & \underline{Q}^t \end{array}\right|$$

In the limit (analogously to the sum of a geometric progression) we have

$$\mathrm{Lim}_{t\to\infty} \sum_{i-0}^{t-1} \underline{Q}^i\underline{R} = (\underline{I} - \underline{Q})^{-1}\underline{R}$$

The ***fundamental matrix*** is defined as $\mathbf{F} = (\mathbf{I} - \mathbf{Q})^{-1}$

So, we have, approximately, as t gets very large,

$$\underline{P}^t = \left|\begin{array}{c:c} \underline{I} & \underline{O} \\ \hdashline \underline{F}\,\underline{R} & \underline{Q}^t \end{array}\right|$$

We can see that **F R** = probability of eventually being absorbed, as shown from our previous "accounts" example:

$$\underline{F}^{-1} = (\underline{I} - \underline{Q}) = \left|\begin{array}{ccc} 1 & -0.4 & 0 \\ 0 & 1 & -0.25 \\ 0 & 0 & 1 \end{array}\right|$$

We **invert** this (see 1.9.1 below) to find matrix F

$$\underline{F} = (\underline{I} - \underline{Q})^{-1} = \begin{vmatrix} 1 & 0.4 & 0.1 \\ 0 & 1 & 0.25 \\ 0 & 0 & 1 \end{vmatrix}$$

Hence,

$$\underline{F}\,\underline{R} = \begin{vmatrix} 1 & 0.4 & 0.1 \\ 0 & 1 & 0.25 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} 0.6 & 0 \\ 0.75 & 0 \\ 0.5 & 0.5 \end{vmatrix} = \begin{vmatrix} 0.95 & .05 \\ 0.875 & 0.125 \\ 0.5 & 0.5 \end{vmatrix}$$

For example, if an invoice is in the second month of being overdue, the Probability that it will eventually be paid = 0.875 and the probability that it will eventually be written off as a bad debt = 0.125.

1.9.1 Inverting a matrix

There are various ways of **calculating the inverse of a given matrix** of which the following is fairly simple. [*Of course, for all but small matrices the calculation would be done by computer.*]

Put the given matrix and the unit matrix (**I**) "side by side" and reduce the given matrix to **I** by elementary row operations, and then the inverse will be formed in the original **I** location.

In the following example the current ***pivot*** element is "bolded" for clarity. Also, as the first column is already in **I** form we start with the second column.

$$\left|\begin{array}{ccc|ccc} 1 & -0.4 & 0 & 1 & 0 & 0 \\ 0 & \mathbf{1} & -0.25 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right| =$$

$$\left|\begin{array}{ccc|ccc} 1 & 0 & -0.1 & 1 & 0.4 & 0 \\ 0 & 1 & -0.25 & 0 & 1 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 1 \end{array}\right| =$$

$$\left|\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0.4 & 0.1 \\ 0 & 1 & 0 & 0 & 1 & 0.25 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right| =$$

1.10 Definitions: Transient states; Cycling; Recurrent state

Transient State: State i is transient if there is a state j that is reachable from i but i is not reachable from j. For example, in the accounts problem states *1, 2,* and *3* are transient. At the steady state $\pi_i = 0$ for transient states.

Cycling (periodic) Process: This occurs when there are all zeros in ***retention*** cells (diagonal top left to bottom right) and zeros or ones in ***non-retention*** cells. For example,

A → B → C → A → ....
No steady state.

	A	B	C
A	0	1	0
B	0	0	1
C	1	0	0

Recurrent Set: When a process is trapped in a set of states.

For example,

$$
\begin{array}{c|cccc|}
 & A & B & C & D \\
A & 0.75 & 0.25 & 0 & 0 \\
B & 0.50 & 0.50 & 0 & 0 \\
C & 0 & 0.30 & 0.35 & 0.35 \\
D & 0 & 0 & 0.20 & .80
\end{array}
$$

Once *A* or *B* is reached, the system stays "trapped" in the recurrent set {A, B}.

1.11 Recap & wider context

This section (mainly based on the start of Chapter XV of (Feller, 1968)

Reviews and summarises some Markov Chain definitions and terminology

Provides some general examples and application areas.

Note: Markov chains are useful in analysing networks of queues and in evaluating the performance of computer systems. See for example books in DCU library such as (Bolch, 2006). However, we do not explore these aspects here.

1.11.1 Definition of a Markov Chain (1/3)

Independent trials in probability theory: A set of possible outcomes E_1, E_2, …, E_k, … is given, and with each there is an associated probability p_k. The probabilities of sample sequences are defined by the multiplicative property –

- for example, if the possible outcomes of tossing a coin are H and T, then the probability of the sequence TTH is $p_T p_T p_h = (1/2)^3 = 1/8$ (for a fair coin).

The theory of Markov chains considers the simplest generalization of the above "Independent trials in …", which consists in permitting the outcome of any trial to depend on the outcome of the directly preceding trial (and only on it).

- The outcome E_k is no longer associated with a fixed probability p_k, but to every pair (E_j, E_k) there corresponds a conditional probability p_{jk}; given that E_j has occurred at some trial, the probability of E_k at the next trial is P_{jk}.

In addition to the conditional (transition) probabilities p_{jk}, we must also be given

- The probability a_k of the outcome E_k at the initial trial.

It follows that the probabilities of sample sequences corresponding to 2, 3 or 4 trials etc must be defined by

- $P\{(E_j, E_k)\} = a_j p_{jk}$, $P\{(E_j, E_k, E_r)\} = a_j p_{jk} p_{kr}$, $P\{(E_j, E_k, E_r, E_s)\} = a_j p_{jk} p_{kr} p_{rs}$ etc

To summarize, a sequence of trials with possible outcomes E_1, E_2, … is called a Markov chain if the probabilities of sample sequences are defined (as above) in terms of a probability distribution $\{a_k\}$ for E_k at the initial (zeroth) trial and fixed conditional probabilities p_{jk} of E_k given that E_j has occurred at the preceding trial.

For applications we had a slightly different terminology:

- The possible outcomes E_k are called the possible states of the system
- Instead of saying that the nth trial results in E_k we say that the nth step leads to E_k or that E_k is entered at the nth step
- p_{jk} is called the probability of a transition from E_j to E_k

We imagine the trials are performed at a uniform rate so that the number of the step serves as "time" parameter.

As before, the transition probabilities are arranged in a matrix whose row sums are one:

$$\begin{vmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}$$

Note: There may be a finite or infinite number of states.

1.11.2 More examples

Example A: When there are only 2 possible states E_1 and E_2, P has the form

$$\begin{vmatrix} 1-p & p \\ \alpha & 1-\alpha \end{vmatrix}$$

For example, a particle moves along the x-axis so that its absolute speed is constant but the direction can be reversed. States E_1 and E_2 correspond to moving positively or negatively, respectively. p and α are the probabilities of a reversal when particle is moving right and left, respectively.

Exercise: Calculate the corresponding steady-state.

Example B (Random walk with absorbing barriers): Let the possible states be $\{E_0, E_1, \ldots, E_r\}$ and consider the transition matrix:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

From each of the "interior" states $E_1, \ldots, E_{r-1}$ transitions are possible to the right and left neighbours (with $p_{i,i+1} = p$ and $p_{i,i-1} = q$). However, no transition is possible from either E_0 or E_r to any other states – they are **absorbing states**. This can be regarded as a model of a random walk with absorbing barriers at 0 and r.

Possible initial distributions $\{a_k\}$ for this system could be (i) $a_z = 1$ (start from E_z) or (ii) $a_k=1/(r+1)$ – randomly chosen distribution.

Variants of this model are possible (e.g. reflecting barriers or cyclical walks).

<u>*A specific question for Example B*</u>:

(i) Determine the fundamental matrix for the case n = 5. You should make use of the note below in your answer.

(ii) For the case n = 5, find the long run probabilities of each internal state transitioning to either E_0 or E_n. What are the values of these probabilities in the specific case of p = q = 1/2 ?

Note:

$$\begin{bmatrix} 1 & -p & 0 & 0 \\ -q & 1 & -p & 0 \\ 0 & -q & 1 & -p \\ 0 & 0 & -q & 1 \end{bmatrix}^{-1} = \frac{1}{1-3pq+p^2q^2} \begin{bmatrix} 1-2pq & p-p^2q & p^2 & p^3 \\ q-pq^2 & 1-pq & p & p^2 \\ q^2 & q & 1-pq & p-p^2q \\ q^3 & q^2 & q-pq^2 & 1-2pq \end{bmatrix}$$

<u>*Solution to question:*</u>

(i) For n = 5,

$$\mathbf{P} = \begin{vmatrix} 1\ 0\ 0\ 0\ 0\ 0 \\ q\ 0\ p\ 0\ 0\ 0 \\ 0\ q\ 0\ p\ 0\ 0 \\ 0\ 0\ q\ 0\ p\ 0 \\ 0\ 0\ 0\ q\ 0\ p \\ 0\ 0\ 0\ 0\ 0\ 1 \end{vmatrix}$$

Next, re-arrange to put absorbing states together:

$$\left|\begin{array}{c:c} 1\ 0 & 0\ 0\ 0\ 0 \\ 0\ 1 & 0\ 0\ 0\ 0 \\ \hdashline q\ 0 & 0\ p\ 0\ 0 \\ 0\ 0 & q\ 0\ p\ 0 \\ 0\ 0 & 0\ q\ 0\ p \\ 0\ p & 0\ 0\ q\ 0 \end{array}\right|$$

or

$$\left|\begin{array}{c:c} \mathbf{I}_{2x2} & \mathbf{0}_{2x3} \\ \hdashline \mathbf{R}_{4x2} & \mathbf{Q}_{4x4} \end{array}\right|$$

The fundamental matrix $\mathbf{F} = (\mathbf{I} - \mathbf{Q})^{-1}$ =

$$\begin{vmatrix} 1 & -p & 0 & 0 \\ -q & 1 & -p & 0 \\ 0 & -q & 1 & -p \\ 0 & 0 & -q & 1 \end{vmatrix}^{-1}$$

Using the given note, it follows that

$$\mathbf{F} = 1/(1\text{-}3pq+p^2q^2) \begin{vmatrix} 1\text{-}2pq & p\text{-}p^2q & p^2 & p^3 \\ q\text{-}pq^2 & 1\text{-}pq & p & p^2 \\ q^2 & q & 1\text{-}pq & p\text{-}p^2q \\ q^3 & q^2 & q\text{-}pq^2 & 1\text{-}2pq \end{vmatrix}$$

(ii) The long run probabilities of each internal state transitioning to either E_0 or E_n are given by **FR** = (check)

$$\mathbf{FR} = 1/(1-3pq+p^2q^2)\left|\begin{matrix} q-2pq^2 & p^4 \\ q^2-pq^3 & p^3 \\ q^3 & p^2-p^3q \\ q^4 & p-2p^2q \end{matrix}\right|$$

For the particular case of p = q = ½, we have

$$\mathbf{FR} = (16/5)\left|\begin{matrix} 1/4 & 1/16 \\ 3/16 & 1/8 \\ 1/8 & 3/16 \\ 1/16 & 1/4 \end{matrix}\right|$$

Finally, we have

$$\begin{array}{c c} & \begin{matrix} E_0 & E_5 \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{matrix} & \left|\begin{matrix} 4/5 & 1/5 \\ 3/5 & 2/5 \\ 2/5 & 3/5 \\ 1/5 & 4/5 \end{matrix}\right| \end{array}$$

For example, there is a 40% probability of going, in the long run, from state E_3 to E_0 and a 60% probability of going from E_3 to E_5.

Example C (An example from Physics – Ehrenfist model of diffusion): Similar to the last example we consider a chain with the ρ+1 states E_0, E_1, …, E_ρ and transitions possible only to right and left neighbours. This time we put $p_{j,j+1}$ = 1-j/ρ and $p_{j,j-1}$ = j/ρ so that the matrix is:

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \rho^{-1} & 0 & 1-\rho^{-1} & 0 & \cdots & 0 & 0 \\ 0 & 2\rho^{-1} & 0 & 1-2\rho^{-1} & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 0 & \rho^{-1} \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

Note 1: Feller goes on to present a different diffusion model called Bernoulli-Laplace.

Note 2: There are several other examples discussed by Feller including

- from cell and population genetics
- from study of recurrent events.

In one example of recurrent events, a system is considered which at a given stage has a probability

- of aging (by say a year) or else
- being set back to age 0 ("rejuvenated")

Of interest in such systems is characterising the ***recurrence time*** of returning to 0.

Note 3: Both examples B and C have (multiple) absorbing events so that it would be possible to compute corresponding **fundamental matrices**.

Exercise: (i) Calculate the fundamental matrix for Example B when r=4 (for arbitrary p and for p=1/2). Similarly for Example C with $\rho = 4$.

Example D: (from (Montgomery, 2009))

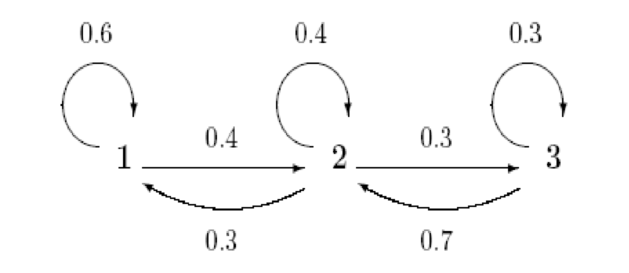
"Sociologists have long been interested in social mobility – the transition of individuals between social classes defined on the basis of income or occupation. Some research has focused on intergenerational mobility from parent's class to child's class, while other research has examined intra-generational mobility over an individual's life course. … here, we'll develop a simple hypothetical example.

Consider a society with three social classes. Each individual may belong to the lower class (state 1), the middle class (state 2), or the upper class (state 3). Thus, the social class occupied by an individual in generation t may be denoted by $s_t \in \{1, 2, 3\}$. Further suppose that each individual in generation t has exactly one child in generation t+1, who has exactly one child in generation t+2, and so on."

"Finally, suppose that intergenerational mobility is characterized by a (3×3) transition matrix which does not change over time. Under these conditions, a single "family history" – the sequence of social classes $(s_0, s_1, s_2, \ldots)$ – is a Markov chain.

To offer a numerical example, suppose that intergenerational mobility is described by the transition matrix

$$P = \begin{bmatrix} 0.6 & 0.4 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0 & 0.7 & 0.3 \end{bmatrix}$$



"Thus, a child with a lower-class parent has a 60% chance of remaining in the lower class, has a 40% chance to rise to the middle class, and has no chance to reach the upper class. A child with a middle-class parent has a 30% chance of falling to the lower class, a 40% chance of remaining middle class, and a 40% chance of rising to the upper class. Finally, a child with an upper-class parent has no chance of falling to the lower class, has a 70% chance of falling to the middle class, and has a 30% chance of remaining in the upper class."

Exercise: Show that the steady-state (long-run) distribution is {21/61, 28/61, 12/61} or (approximately) {0.3443, 0.4590, 0.1967}.

A "macro-level" interpretation

"… we have so far maintained a "**micro-level**" interpretation of Markov chain processes. For instance, in the social mobility example, we adopted the perspective of a particular individual, and considered how that individual's social class affected the "life chances" of his child and grandchild and subsequent descendants.

Thus, the initial condition x_0 reflected the individual's class, and the probability vector $\mathbf{x}_t$ reflected the probability distribution over classes for a particular descendent in generation t.

But it is also possible to adopt a "**macro-level**" interpretation of this process. Given a large population in which each individual belongs to one of the social classes, we may interpret $\mathbf{x_0}$(i) as the share of the population in class i in generation 0. For instance, setting $\mathbf{x}_0$ = [0.2 0.3 0.5] , we are assuming that 20% of the population belongs to the lower class, that 30% belong to the middle class, and that 50% belong to the upper class"

1.12 Brief note on other Eigenvalue & Markov Chain applications

The set of slides "5. Matrix Models" (Jung, n.d.) are on Loop. Here we just highlight a few elements from these slides to give some additional insights on Markov Chains and on Eigenvalue Problem applications generally.

5.1 Dynamical Systems and Markov Chains:

Slides 4-6 present an example such as we have had before:

> "Suppose that two competing television news channels, channel 1 and channel 2, each have 50% of the viewer market at some initial point in time. Assume that over each one-year period channel 1 captures 10% of channel 2's share,

and channel 2 captures 20% of channel 1's share. What is each channel's market share after one year?"

There is another such example (slides 9-10) regarding the migration pattern of a tagged lion!

The rest of section 5.1 (as far as slide 15) summarises (what we had before) about steady-state and how to calculate it. It contains examples.

5.2 Leontief Input-Output Models (16-25):

This discusses matrix models of an **open** economy.

In fact, of more direct interest as an eigenvalue problem is a **closed economy** in which all the industries are inter-related such that all their outputs are used as inputs by themselves. This leads to a problem of the form

$$\mathbf{A}_{nxn}\mathbf{p}_{nx1} = \mathbf{p}_{nx1}$$

where n is the number of industries, p_j is the price charged by industry j for its total output, and **A** is called the **consumption matrix** (gives the fraction of each industry's output consumed by the industries in the economy).

In an **open** economy some output is left over for consumption by non-industry consumers.

5.3 Gauss-Seidel and Jacobi Iteration (26-37):

This presents two methods for solving sparse systems of linear equations. Not relevant to our purpose here, more a topic for numerical analysis.

5.4 The Power Method (38-60)

"*In this section we will discuss an algorithm that can be used to approximate the* ***eigenvalue with greatest absolute value*** *and a* ***corresponding eigenvector***."

- The details of the "power method" algorithm would be discussed in a numerical analysis class and are not our concern here.

- However, the application presented (53-59), "An Application of The Power Method to Internet Searches", is interesting; see particularly pages 58 and 59 for some specific calculations.

Note: There is a description of the "The Google PageRank Algorithm" in section of (Hastie, et al., 2008), including reference to Markov Chain theory in that *PageRank* may be viewed as a Markov Chain. The discussion includes the text:

"*The PageRank algorithm considers a webpage to be important if many other webpages point to it. However the linking webpages that point to a given page are not treated equally: the algorithm also takes into account both the importance (PageRank) of the linking pages and the number of outgoing links that they have.*

Linking pages with higher PageRank are given more weight, while pages with more outgoing links are given less weight. …".

1.13 Problems

Question 1

A commercial copy centre uses a large photocopier that deteriorates rather rapidly, in terms of quality of copies produced, under heavy usage. The machine is examined at the end of each day to determine its status with respect to copy quality. The inspection results are classified as follows:

State	***Condition of photocopier***
1	Excellent
2	Acceptable – some deterioration in quality
3	Marginal – significant deterioration in quality
4	Unacceptable quality –repairs required

Based on past data collected concerning the operation of the photocopier, the condition of the machine over a period of time has been described by the transition matrix given below:

		State on the following day			
		1	**2**	**3**	**4**
	1	0	.8	.1	.1
State on	**2**	0	.6	.2	.2
one day	**3**	0	0	.5	.5
	4	1	0	0	0

The expected costs associated with the various states of the machine are as follows:

State	***Expected cost per day***
1	€0
2	€100 (returning copies)
3	€400 (refinds to dissatisfied customers)
4	€800 (repair cost and downtime)

(i) What is the expected cost associated with this policy of repairing the machine to a like-new condition each time it reaches state 4?

(ii) The company is implementing a new policy whereby when the copy quality is found to be marginal, a minor repair is carried out, at a cost of €200, whereby the copy quality is returned to acceptable (state 2). Would it be worthwhile implementing this policy?

Solution Hints

(i) Find the steady state steady state proportions. [**Ans**: $\pi_1 = 0.2 = \pi_3 = \pi_4$ & $\pi_2 = 0.4$] Then deduce the expected cost. [**Ans**: €280]

(ii) Set up an altered transition matrix and deduce that, for this matrix, $\pi_1 = \pi_3 = \pi_4 = 0.13$ & $\pi_2 = 0.6$. Hence, find the expected cost [**Ans**: €190]

Question 2

A car hire firm has offices at three Irish airports, Dublin, Shannon and Cork. The average weekly demand for cars at the airports is 500 in Dublin, 200 in Shannon and 300 in Cork. The cars may be returned by the hirer at the end of the week to any of these three airports after use, and this occurs with the following probabilities:

		Returned to	
Hired from	**Dublin**	**Shannon**	**Cork**
Dublin	0.3	0.5	0.2
Shannon	0.6	0.2	0.2
Cork	0.4	0.1	0.5

The firm must arrange for its employees to deliver the returned cars between airports so as to match the average demand. The cost is €100 per car moved between Cork and Dublin, or between Shannon and Dublin, and €75 per car moved between Shannon and Cork.

(i) Find the average weekly cost of the delivery service.

(ii) Suppose the firm decides to drop the delivery service and to increase advertising so that all cars available in each airport would be taken up. Find the long term distribution of cars at the airports.

Solution Hints:
(i) Remember that in each week, to meet the demand, it must be arranged to have 500 in Dublin, 200 in Shannon and 300 in Cork.500. [**Ans**: Total cost = €11,750]
(ii) Long term distribution of cars at airports is

Dublin	418
Shannon	297
Cork	286

1.14 Some worked examples

Example 1: A simple three state model of the weather (Precipitation (P), Cloudy (C) and Sunny (S)) has been formulated, and the following (stationary) state transition probabilities have been observed for a particular geographic region:

		Weather now		
		P	**C**	**S**
Weather in previous hour	**P**	0.4	0.3	0.3
	C	0.2	0.6	0.2
	S	0.1	0.1	0.8

For example, if there was precipitation one day ago there is a 40% chance that there is precipitation now.
(i) Explain why a steady-state exists for this example.
(ii) Calculate the steady-state probabilities.
(iii) What is the probability that the weather for eight consecutive days is "Sunny-Sunny-Sunny-Precipitation-Precipitation-Sunny-Cloudy-Sunny"?

Solution:

(i) Transition matrix **P** is clearly irreducible as all of $\mathbf{P}^1$ elements are greater than 0. Hence, a steady-state exists. (Refer to Section 1.6.2)

(ii) Let $[\pi_1, \pi_2, \pi_3]$ be the steady-state probabilities so that

$$[\pi_1, \pi_2, \pi_3] = [\pi_1, \pi_2, \pi_3] \begin{vmatrix} .4 & .3 & .3 \\ .2 & .6 & .2 \\ .1 & .1 & .8 \end{vmatrix}$$

=>

$\pi_1 = 0.4\pi_1 + 0.2\pi_2 + 0.1\pi_3$ (a)

$\pi_2 = 0.3\pi_1 + 0.6\pi_2 + 0.1\pi_3$ (b)

$\pi_3 = 0.3\pi_1 + 0.2\pi_2 + 0.8\pi_3$ (c)

(a) & (b)=>

$\pi_1 - \pi_2 = 0.1\pi_1 - 0.4\pi_2 \Rightarrow \pi_1 = (2/3)\pi_2$

(b) & (c) =>

$\pi_2 - \pi_3 = 0.4\pi_2 - 0.7\pi_3 \Rightarrow \pi_3 = 2\pi_2$

Also, $1 = \pi_1 + \pi_2 + \pi_3 \Rightarrow 1 = (2/3)\pi_2 + \pi_2 + 2\pi_2 \Rightarrow \pi_2 = 3/11$

Hence, the steady-state proportions or probabilities are

$[\pi_1, \pi_2, \pi_3] = [2/11, 3/11, 6/11]$

(iii) The required sequence requires that it is Sunny at the outset (π_3) followed by 7 transitions:

$(\pi_3)P_{33}\ P_{33}\ P_{31}\ P_{11}\ P_{13}\ P_{32}\ P_{23} = 0.00084 = 0.84\text{x}10^{-4}$

Example 2: The following states have been identified for final testing of a software product:

A. First test set completed
B. Second test completed
C. Third test set completed
D. Released
E. Discarded

There is 0.4 probability of going from A to B and 0.6 probability of going from A to D. There is 0.5 probability of going from B to C and 0.5 probability of going from B to D. The probability of going from C to D is 0.3 and from C to E is 0.7. If state D is entered then the probability is 1 of remaining in that state, and similarly for state E.

(i) Write out the matrix of transition probabilities and explain why a steady-state does not exist in this case.

(ii) Calculate the ***fundamental matrix*** for this system.

(iii) What are the probabilities that the software product will be released eventually in each of the following cases:

the first test has been completed
the second test has been completed
the third test has been completed?

Solution:

(i)

	A	B	C	D	E
A	0	.4	0	.6	0
B	0	0	0.5	0.5	0
C	0	0	0	.3	.7
D	0	0	0	1	0
E	0	0	0	0	1

A steady-state does not exist because there are 2 absorbing states D and E (see section 1.6.2).

(ii) Re-arrange the matrix (for convenience) to

$$\begin{array}{c|ccccc|} & D & E & A & B & C \\ D & 1 & 0 & 0 & 0 & 0 \\ E & 0 & 1 & 0 & 0 & 0 \\ A & 0.6 & 0 & 0 & .4 & 0 \\ B & 0.5 & 0 & 0 & 0 & 0.5 \\ C & 0.3 & 0.7 & 0 & 0 & 0 \end{array}$$

This is of the form

$$\mathbf{P} = \left|\begin{array}{c:c} \mathbf{I} & \mathbf{O} \\ \hdashline \mathbf{R} & Q \end{array}\right|$$

If F is the fundamental matrix then $\mathbf{F}^{-1} = \mathbf{I}\text{-}\mathbf{Q}$ i.e.

$$\mathbf{F}^{-1} = \begin{vmatrix} 1 & -.4 & 0 \\ 0 & 1 & -.5 \\ 0 & 0 & 1 \end{vmatrix}$$

This needs to be inverted to get **F**. One could probably guess **F**, in fact, but the following systematic procedure works (see section 1.9.1):

$$\left|\begin{array}{ccc|ccc} 1 & -.4 & 0 & 1 & 0 & 0 \\ 0 & 1 & -.5 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right|$$

->

$$\left|\begin{array}{ccc|ccc} 1 & 0 & -.2 & 1 & .4 & 0 \\ 0 & 1 & -.5 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right|$$

->

$$\left|\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & .4 & .2 \\ 0 & 1 & 0 & 0 & 1 & .5 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right|$$

Hence, **F** =

$$\begin{vmatrix} 1 & .4 & .2 \\ 0 & 1 & .5 \\ 0 & 0 & 1 \end{vmatrix}$$

(iii) Calculate **FR** =

$$\begin{vmatrix} 1 & .4 & .2 \\ 0 & 1 & .5 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} .6 & 0 \\ .5 & 0 \\ .3 & .7 \end{vmatrix}$$

=

$$\begin{vmatrix} .86 & .14 \\ .65 & .35 \\ .3 & .7 \end{vmatrix}$$

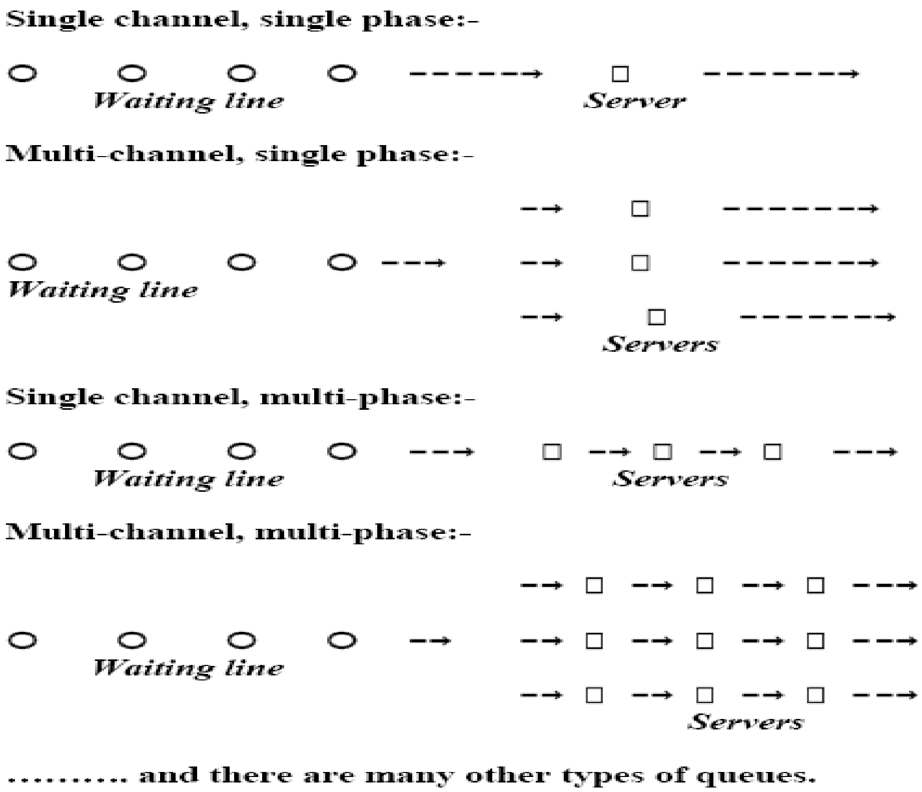
The first column gives the required probabilities.

2. Queue Theory

2.1 Introduction

This topic is covered in general Operations Research books such as (Taha, 2007) as well as more specialised or advanced texts such as (Feller, 1968).

We start with a depiction of different kinds of queue systems although we will cover only rather simple systems in detail.



There are various **measures of performance** (or **operating characteristics**) that are relevant in analysing queue systems:

P_n = probability of n items in the system (queue + service)

L = Mean number of items in the system

L_q = Mean number of items in the queue

W = Mean time spent in the system

W_q = Mean time spent in the queue

There are two **probability distributions** that are very important for queue theory, namely

<u>Poisson distribution</u> (especially to model arrivals rate)

$P(r) = (e^{-\lambda} \lambda^r) / r!$

where

r = arrivals/(time unit)

λ = (mean number of arrivals)/(time unit)

Negative Exponential distribution (e.g. service times and inter-arrival times[1]))

$f(t) = \mu e^{-\mu t}$

where

t = service time

μ = mean service time

Assumptions are invariably made to simplify the analysis of queue systems. Assumptions commonly made include:

- Arrivals rate distribution is Poisson with mean of λ
- Service times distribution is Negative Exponential with mean = $1/\mu$ (i.e. Service rate distribution is Poisson)
- Queue discipline is *FIFO (first in, first out)*
- No reneging (customers stay in queue)
- No balking (baulking)
- Mean service rate > mean arrival rate ($\mu > \lambda$)
- Calling population is infinite (∞)
- Capacity of system is unlimited
- Steady-state characteristics – queue has settled down. (contrast with Transient characteristics such as at rush hour – see example below).

A shorthand **notation due to Kendal** is commonly used for queues.

a/b/c : d/e/f

where

a = arrivals distribution

b = service distribution

c = # of servers

d = queue discipline

e = system capacity

[1] If the arrival rate is Poisson distributed then, from Probability Theory, the inter-arrival times are distributed according to the negative exponential distribution.

f = size of calling population

Distributions that commonly arise in Queue theory are identified as follows in Kendal's notation:

$M \rightarrow$ Poisson

$E_k \rightarrow$ Erlang

$G \rightarrow$ General

$D \rightarrow$ Deterministic

Note: We do not have a symbol for negative exponential as it is the counter part of Poisson (as explained).

To give a feeling for the **difference between transient and steady-state queues** we present the following:

Example of a transient 2-state system

Possible interpretation :
A = No customers present
B = One customer being served
Note: Customers who arrive while system is in state B leave without being served & do not return.

Arrival rate & service rate (e.g. per hour) :

```
In[22]:=  λ = 3.5; μ = 4;
```

Specify initial probabilities of being in state A and state B::

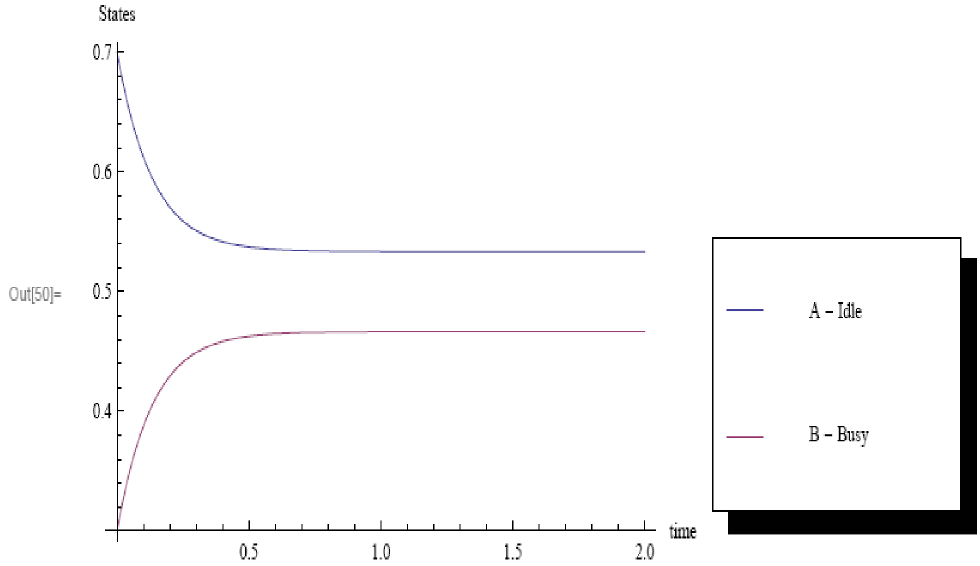
```
In[23]:=  pA0 = 0.7; pB0 = 0.3;
```

Define transient solution (solution of a differential equation - see **NOTE** below):

```
In[24]:=  pA[t_] := μ / (λ + μ) * (1 - Exp[- (λ + μ) * t]) + pA0 * Exp[- (λ + μ) * t]

In[25]:=  pB[t_] := λ / (λ + μ) * (1 - Exp[- (λ + μ) * t]) + pB0 * Exp[- (λ + μ) * t]
```

Following is a graph of these transient (time-dependent) quantities:



Note on derivation of transient (differential equations) and steady-state (difference equations) solutions:

(1) The transient solution is the solution of the differential equation
Here, prime (') means "derivative with respect to t":

```
pA'[t] = - λ pA[t] + μ pB[t]

pB'[t] =   λ pA[t] - μ pB[t]
```

(2) The steady solution is the solution of the difference equation :

```
0 = - λ pA + μ pB and 0 =  λ pA - μ pB
```

For both transient and steady-state we must satisfy the "probability" condition that pA + pB = 1

So, for the steady - state, we have pB = (λ/μ) pA

2.2 Little's relationship (for any queue system)

Little's formulae holds for **all** queuing systems at the steady state:-

$L = \lambda W$ (e.g. $\lambda = 4$ per hour, $W = ½$ hour then $L = 2$)

and

$L_q = \lambda W_q$

We also have for all queue systems in the steady state that

Expected time in system (W) = Expected time in queue + expected service time.

That is:

$W = W_q + 1/\mu$

From this it follows, using Little's relationship, that

$L = L_q + \lambda/\mu$

2.3 M/M/1 Queue (Steady-State)

We derive an equation set for an M/M/1 queue, that is a queue with random arrivals and services.

Define an interval of time Δt whose size is such that there is at most 1 arrival and at most 1 departure during Δt. *(Essentially, this means we ignore terms in $(\Delta t)^2$)*

Let λ be the mean arrival rate (*e.g. $\lambda = 4$ would mean there are 4 arrivals in unit time which is equivalent to saying that the average time between arrivals is 0.25 time units*)

Let μ be the mean service rate (*e.g. $\mu = 5$ would mean there are (potentially) 5 departures in unit time which is equivalent to saying that the average service time is 0.20 time units*)

We have the following probabilities:

Prob of 1 arrival in $\Delta t = \lambda\Delta t$ (*e.g. if* $\lambda = 4$ and $\Delta t = 0.1$ *time units then* $\lambda\Delta t = 0.4$)

<u>Hence</u>, prob of no arrival in $\Delta t = 1 - \lambda\Delta t$

Prob of 1 departure in $\Delta t = \mu\Delta t$ (*e.g. if* $\mu = 5$ and $\Delta t = 0.1$ *time units then* $\mu\Delta t = 0.5$)

<u>Hence</u>, prob of no departure in $\Delta t = 1 - \mu\,\Delta t$

Prob of 1 arrival <u>and</u> 1 departure $= (\lambda\Delta t)\,.\,(\mu\Delta t) = \lambda\mu(\Delta t)^2 \approx 0$

Prob of no arrival <u>&</u> no departure $= (1-\lambda\Delta t)(1-\mu\Delta t) = 1-\lambda\Delta t-\mu\Delta t+\lambda\mu(\Delta t)^2 \approx 1-\lambda\Delta t-\mu\Delta t$

Let $P_n(t+\Delta t)$ = probability that there are n units in the system at time $t + \Delta t$.

We have

$P_0(t+\Delta t)$ = (Prob of no arrival in time Δt)(Prob of 0 in system at time t)

+ (Prob of 1 departure in time Δt)(Prob of 1 in system at time t)

$\Rightarrow P_0(t+\Delta t) = (1-\lambda\Delta t)P_0(t) + (\mu\Delta t)P_1(t)$ (1)

Similarly,

$P_1(t+\Delta t)$ = (Prob of 1 arrival in time Δt)(Prob of 0 in system at time t)

+ (Prob of 1 arrival & 1 departure in time Δt) (Prob of 1 in system at time t)

+ (Prob of 1 departure in time Δt)(Prob of 2 in system at time t)

=> $P_1(t+\Delta t) = (\lambda\Delta t)P_0(t) + (1 - \lambda \Delta t - \mu \Delta t)P_1(t) + (\mu \Delta t)P_2(t)$ (2)

In general, for $n > 0$,

$P_n(t+\Delta t) = (\lambda \Delta t)P_{n-1}(t) + (1 - \lambda \Delta t - \mu \Delta t)P_n(t) + (\mu \Delta t)P_{n+1}(t)$ (3)

At the steady state $P_n(t+\Delta t) = P_n(t) = P_n$ (that is, no longer dependent on time)

Hence, (1) and (3) simplify to

$$0 = -\lambda P_0 + \mu P_1 \qquad (1')$$

$$0 = \lambda P_{n-1} - (\lambda + \mu)P_n + \mu P_{n-1} \qquad (3')$$

From (1'), solving for P_1 :-

$$P_1 = P_0 (\lambda/\mu) \qquad (4)$$

From (3'), solving for P_{n+1} :

$$P_{n+1} = ((\lambda + \mu)/\mu) P_n - (\lambda /\mu) P_{n-1} \qquad (5)$$

e.g. $P_2 = ((\lambda + \mu)/\mu)P_1 - (\lambda/\mu)P_0 = (\lambda/\mu)^2P_0$

In general, it follows that

$$P_n = (\lambda/\mu)^nP_0 \quad (n = 1, 2, 3, \ldots) \qquad (6)$$

Note: The sum of an infinite geometric sequence a, ar, ar^2, ar^3, … ($|r| < 1$) is $a/(1-r)$.

(1) Finding P_0 (proportion of idle time of queue):

By definition the sum of the probabilities is 1 that is $1 = \Sigma P_n$

Using (6) we have $1 = P_0 \Sigma(\lambda /\mu)^n$

By assumption, $\lambda /\mu < 1$, so $1 = P_0 /(1 - \lambda /\mu)$ (using note on a GP)

Hence, $\boxed{P_0 = 1 - \lambda /\mu}$

Therefore, substituting into (6), $\boxed{P_n = (1 - \lambda /\mu) (\lambda /\mu)^n}$

(2) Finding L (average number of items in the system): We let $\rho = \lambda/\mu$

$$\begin{aligned} L &= \Sigma_{n=0}\, n.P_n = \Sigma_{n=1}\, n.P_n = \Sigma_{n=1}\, n.(1 - \lambda /\mu)(\lambda /\mu)^n \\ &= (1 - \rho)\Sigma_{n=1}\, n\rho^n = (1 - \rho)\rho\Sigma_{n=1}\, n\rho^{n-1} = (1 - \rho)\rho\Sigma_{n=1}\, d(\rho^n)/d\rho \\ &= (1 - \rho)\rho d[\Sigma_{n=1}\, \rho^n]/d\rho = (1 - \rho)\rho d[\rho/(1-\rho)]/d\rho \\ &= (1 - \rho)\rho/(1-\rho)^2 = \rho/(1-\rho) = (\lambda/\mu) / (1- \lambda /\mu) \end{aligned}$$

Hence, $\boxed{L = \lambda /(\mu - \lambda)}$

Summary of formulae for steady-state M/M/1:

$P_0 = 1 - \lambda/\mu$
$P_n = (1 - \lambda/\mu)(\lambda/\mu)^n$
$L = \lambda/(\mu - \lambda)$

Using Little's relationship and related formulae ($L = \lambda W$, $L_q = \lambda W_q$, $W = W_q + 1/\mu$, $L = L_q + \lambda/\mu$) we also have for steady-state M/M/1:

$W = L/\lambda = 1/(\mu - \lambda)$
$W = W_q + 1/\mu \quad \Rightarrow W_q = \lambda/(\mu(\mu - \lambda))$
$L_q = \lambda W_q = \lambda^2/(\mu(\mu - \lambda))$

Example: Customers arrive at an ATM machine at an average rate of 20 per hour (assume the arrivals rate is described by a Poisson distribution). The amount of time they spend at the machine takes on average two minutes, but can vary from customer to customer (assume negative exponential distribution). Calculate the average number of customers in the system, the average time they spend in the system, the average number in the queue, the average time spent in the queue and the percentage of the time the machine is idle.

Solution: $\lambda = 20$, $\mu = 30$

Average no. of customers in system = L = 20/(30-20) = 2

Average time in system = W = 2/20 = 0.1 hours = 6 minutes.

Average no. in queue = $L_q = 20^2/(30(30-20)) = 400/300 = 4/3$

Average time spent in queue = $W_q = W - 1/\mu = 0.1 - 1/30 = 1/15$ hours = 4 minutes

Percentage of idle time = $P_0 = 1 - 20/30 = 1-2/3 = 1/3$ or 33.33%

2.4 The M/M/s Queue (steady-state)

In this case we will (for now, but see section 2.5) just state the formulae without deriving them. ***Remember that Little's formulae still hold***. However, the formula for P_0 will be different (as will those for P_n, L, L_q etc).

We have the following:

s = Number of servers

Service distributions are assumed to be the same for all servers.

μ = average service rate per server

Mean effective service rate = $s\mu$

We often let $\rho = \lambda/(s\mu)$ = traffic intensity.

Note: We must have $s\mu > \lambda \Rightarrow \rho < 1$ (for steady-state to make sense)

We state the formulae (from which all quantities can be derived):

$$P_0 = 1 \Big/ \left[\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!(1-\lambda/s\mu)} \right]$$

$$P_n = ((\lambda/\mu)^n / n!)\, P_0 \qquad \text{(for } n \le s\text{)}$$
$$P_n = ((\lambda/\mu)^n / (s!\, s^{n-s}))\, P_0 \qquad \text{(for } n > s\text{)}$$

$$L_q = \frac{P_o (\lambda/\mu)^s \rho}{s!\,(1-\rho)^2}$$

Example: A computer room has three printers which can each print an average of five jobs per minute. The average number of jobs entering a single queue for the three machines is twelve per minute. Assuming the queue is M/M/3, calculate the following:

(i) The average time a job is in the system

(ii) The average number of jobs in the system

(iii)Whether any time would be saved for the customers if the three slow printers were replaced by one fast printer, working at 15 jobs per minute.

Solution: We will have to first calculate P_0 (as is clear from the formulae above)

$S = 3, \lambda = 12, \mu = 5 \Rightarrow \rho = 12/(3 \times 5) = 4/5 = 0.8$ and $\lambda/\mu = 2.4$

Then, $P_0 = [1 + 2.4 + 2.4^2/2 + 2.4^3/\{6(1-0.8)\}]^{-1} = 0.056$ (***check!***)

We are given the formula for $L_q = 0.056(2.4)^3(0.8)/(6(1-0.8)^2) = 2.58$ (***check!***)

(i) $W = L_q/\lambda + 1/\mu = 2.58/12 + 1/5 = 0.415$ minutes

(ii) $L = \lambda W = 12(0.415) = 4.98$

(iii) Apply M/M/1 with $\lambda = 12$, $\mu = 15 \Rightarrow W = 1/(\mu - \lambda) = 1/3 = 0.333$ minutes. So time would be saved ($0.415 - 0.333 = 0.082$ minutes).

2.5 Some generalizations of M/M/1 & M/M/s

For the M/M/1 system we had the following steady-state equations for P_n:

$$P_1 = P_0 (\lambda/\mu)$$

$$P_{n+1} = ((\lambda + \mu)/\mu)\, P_n - (\lambda/\mu) P_{n-1}$$

Or

$$0 = -\lambda P_0 + \mu P_1$$

$$0 = -(\lambda + \mu)P_n + \lambda P_{n-1} + \mu P_{n+1}$$

where

$\lambda\Delta t$ = probability of an arrival in a time Δt

$\mu\Delta t$ = probability of a departure in a time Δt

From these equations we were able to deduce that

$$P_n = (\lambda/\mu)^n P_0$$

and, using the fact that $\Sigma P_n = 1$, that

$$P_0 = 1 - \lambda/\mu$$

In the following notes, we present a general result (*Birth-And-Death Process, see also section 2.7.4*) of which M/M/1 and M/M/s are special cases, and also present some other special cases.

Probability of n in system (P_n) under general conditions: Our generalisation is to make the probability of arrivals and departures dependent on n. Thus, we let

$\lambda_n\Delta t$ = prob. of an arrival in $[t, t + \Delta t]$ if there is n in system at time t

$\mu_n\Delta t$ = prob. of a departure in $[t, t + \Delta t]$ if there is n in system at time t

Under these conditions, the Steady-State equations are

$$0 = -\lambda_0 P_0 + \mu_1 P_1$$

$$0 = -(\lambda_n + \mu_n)P_n + \lambda_{n-1}P_{n-1} + \mu_{n+1}P_{n+1}$$

From these equations it can be deduced that

$$\boxed{P_n = \frac{\lambda_0\lambda_1\ldots\lambda_n}{\mu_1\mu_2\ldots\mu_{n+1}} P_0} \qquad (*)$$

and, using the fact that $\Sigma P_n = 1$, that

$$\boxed{P_0\{1+\frac{\lambda_0}{\mu_1}+\frac{\lambda_0\lambda_1}{\mu_1\mu_2}\ldots\} = 1 \text{ or } P_0 = \{1+\frac{\lambda_0}{\mu_1}+\frac{\lambda_0\lambda_1}{\mu_1\mu_2}\ldots\}^{-1}} \qquad (**)$$

M/M/1: For this we just have $\lambda_n = \lambda$ and $\mu_n = \mu$

Then, (*) and (**) become

$$P_n = \frac{\lambda\lambda\ldots\lambda}{\mu\mu\ldots\mu} P_0 = \left(\frac{\lambda}{\mu}\right)^n P_0 \qquad (*)\ \text{M/M/1}$$

and

$$P_0\{1+\frac{\lambda}{\mu}+\frac{\lambda\lambda}{\mu\mu}\ldots\} = 1 \text{ or } P_0 = \{1+\frac{\lambda}{\mu}+\frac{\lambda^2}{\mu^2}\ldots\}^{-1} = 1-\frac{\lambda}{\mu} \qquad (**)\ \text{M/M/1}$$

Note: The traffic intensity (also called utilisation factor or load factor) is $\rho = \lambda/\mu$

Immigration-Death process: For this we have $\lambda_n = \lambda$ and $\mu_n = n\mu$ where μ is independent of n. This is interpreted as follows:

whenever a customer arrives a server is made available.

So, when there are n in the system, it means there are n customers each being served. In this case we have

$$P_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n P_0 \qquad (*)\ \text{Immigration-Death}$$

and

$$P_0 = \{1+\frac{\lambda}{1!\,\mu}+\frac{\lambda^2}{2!\,\mu^2}\ldots\}^{-1} = e^{-\lambda/\mu} \qquad (**)\ \text{Immigration-Death}$$

This can be used to model a species, which is born with a constant rate and dies at a rate proportional to the amount (i.e. population) of the species.

M/M/s: This is a more realistic version of the previous one where there is an upper limit (S) on the number of servers. For this we have

$\lambda_n = \lambda$ for all n,

$\mu_n = n\mu$ if $n < S$ but $\mu_n = S\mu$ if $n \geq S$,

where μ is independent of n.

As stated before we must have, for a steady-state to develop, $\rho = \lambda/(S\mu) < 1$. Formulae are a bit "messier" (as we saw in section 2.4):

$$P_n = \frac{\rho^n}{n!} P_0 \text{ if } 1 \le n < S$$

$$P_n = \frac{\rho^n}{S!S^{n-S}} P_0 \text{ if } n \ge S$$

(*) M/M/S

and

$$P_0 = \{1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} \ldots + \frac{\rho^{S-1}}{(S-1)!} + \frac{\rho^S}{S!(1-\frac{\rho}{S})}\}^{-1}$$

(**) M/M/S

Queue with limited waiting room (R):

λ_n = λ if $n < R$,

= 0 if $n \ge R$

μ_n = μ for all n

Calls to telephone exchange with k lines:

λ_n = λ if $n < k$

= 0 if $n \ge k$

μ_n = $n\mu$ for all n ($n \le k$)

There are several other special cases.

2.6 Some formulae for non- M/M systems

2.6.1 Single-Server queues for "customers" all with same priority

M/M/1 Random (Poisson) arrivals

Random (Exponential) service times

M/D/1 Random (Poisson) arrivals

Constant (deterministic) service times

M/G/1 Random (Poisson) arrivals

General service times

Note: Some formulae apply for any queueing discipline of service times but others apply only to first-in, first-out (FIFO) discipline.

Note: M in "M/M/s" stands for "Markovian" or "Memory-less"

Note: In following slides, we are assuming **steady-state** conditions.

	M/M/1	**M/D/1**	**M/G/1**
L	$\rho/(1-\rho)$	$\rho+\rho^2/[2(1-\rho)]$	$\rho + \rho^2/[2(1-\rho)](1+(\mu\sigma_{1/\mu})^2)$

W	$1/[\mu(1-\rho)]$	$(2-\rho)/[2\mu(1-\rho)]$	$1/\mu + \rho/[2\mu(1-\rho)](1+(\mu\sigma_{1/\mu})^2)$
L_q	$\rho^2/(1-\rho)$	$\rho^2/[2(1-\rho)]$	$\rho^2/[2(1-\rho)](1+(\mu\sigma_{1/\mu})^2)$
W_q	$\rho/[\mu(1-\rho)]$	$\rho/[2\mu(1-\rho)]$	$\rho/[2\mu(1-\rho)](1+(\mu\sigma_{1/\mu})^2)$

Notes:

(1) The above set of formulae is not complete.

(2) Assumed given are λ (mean arrival rate), μ (mean service rate) and therefore $\rho = \lambda/\mu$. Also, for general service distributions $\sigma_{1/\mu}$ (standard deviation of service times).

<u>**Example**</u>: In a previous example (ATM machine) we had $\lambda = 20$, $\mu = 30 \Rightarrow \rho = 2/3$. Suppose also that $\sigma_{1/\mu} = 0.5$ minutes (*which may be rather small*) compared to the average serving time of $1/\mu = 1/30$ hours or 2.0 minutes; hence, $1+[\sigma_{1/\mu}/(1/\mu)]^2 = 1 + [0.5/2]^2 = 1.0625$.

We find

	M/M/1	**M/D/1**	**M/G/1**
L	2	4/3 = 1.333	2/3 + (1.0625)2/3 = 1.375
W	1/10 hrs = 6 mins	1/15hrs = 4 mins	1/30 + (1.0625)/30 hrs = 4.125mins
L_q	4/3	2/3 = 0.667	(1.0625)2/3 = 0.708
W_q	1/15 hrs = 4 mins	1/30=hrs = 2 mins	(1.0625)/30 = 0.0354 hrs = 2.125mins

Clearly, the deterministic queue has the best performance measures.

2.6.2 Single-Server queues for "customers" with different priorities but <u>without</u> pre-emption

M/M/1 Random (Poisson) arrivals

Random (Exponential) service times

M/D/1 Random (Poisson) arrivals

Constant service times

M/G/1 Random (Poisson) arrivals

General service times

<u>We do not reproduce the formulae here.</u>

2.6.3 Single-Server queues for "customers" with different priorities & <u>with</u> pre-emption

M/M/1 Random (Poisson) arrivals

Random (Exponential) service times

M/D/1 Random (Poisson) arrivals

Constant service times

M/G/1 Random (Poisson) arrivals

General service times

We do not reproduce the formulae here.

Note: Results are also available for a (steady-state) D/M/1 queue, though not in as convenient a form.

2.6.4 Multiple-Server queues for "customers"

A lot of work has been done in this area though closed form solutions for a general service time distribution have not been found (***check?***).

To get an idea of what has been done you could refer, for example, to (Kleinrock, 1975) (DCU library: 519.82/KLE) *"Presents and develops methods from queueing theory in mathematical language and ... . Step–by–step development of results with careful explanation, and lists of important results ..."* or to a more specific source, such as (Vega, 1998).

2.7 Queues and Markov Processes

2.7.1 Markov Processes; Chapman-Kolmogorov identity

In Section 1.1 we stated that a **stochastic process**, applied to a system with distinct states, means a process in which probabilities are associated with entering a particular state. The probabilities usually depend on the previous history of the system. A **Markov Process** is a particular kind of stochastic process where the probability of entering a particular state depends only on the **last** state occupied …

In Section 1 we discussed **Markov Chains** which, in general involve only countably many states E_1, E_2, … and depend on a discrete time parameter. This means that changes occur only at fixed epochs[2] t = 0, 1, ... For example, we had a newspaper readership application involving "changes" at yearly intervals.

By contrast, if we are dealing with phenomena such as telephone calls or customers arriving at a supermarket checkout, the changes may occur at any time. Mathematically speaking, we shall be concerned with stochastic processes involving only countably many states but depending on a continuous time parameter. For

[2] The term epoch is used to denote points on the time axis.

example, a queue might be in the state of having, at time t, 0 customers or 1 customer or 2 customers and so on.

We introduce the **transition probability** $P_{jk}(t)$ (analogous to that for Markov Chains) meaning the conditional probability of the system being in state E_k at epoch t+s given that at epoch s < t+s the system was in state E_j. We assume a **stationary** or homogeneous process meaning that the probabilities depend only on the duration t but not on position (s) on the time axis.

The basic relationship is the following, called the **Chapman-Kolmogorov identity**,

$$P_{ik}(\tau+t) = \sum_j P_{ij}(\tau)P_{jk}(t)$$

which is based on the following reasoning. It is assumed that $\tau > 0$ and $t > 0$.

> Suppose that at epoch 0 the system is in state E_i. Then, $P_{ij}(\tau)$ is the probability that it will be in state j at epoch τ. Also, if it is state E_j at epoch τ then $P_{jk}(t)$ is the probability that it will be in state k at epoch τ+t.
>
> Therefore, $P_{ij}(\tau)P_{jk}(t)$ is the probability of being in state E_i at epoch 0, at the intermediate state E_j at epoch τ, and finally being at state E_k at epoch τ+t.

The **Chapman-Kolmogorov identity** is just the summation over all possible intermediate states.

A particular case: suppose there are only three possible states E1, E2 and E3. Then, for example,

$$P_{12}(\tau+t) = P_{11}(\tau)P_{12}(t) + P_{12}(\tau)P_{22}(t) + P_{13}(\tau)P_{32}(t)$$

In the following sections, we shall study solutions of the **Chapman-Kolmogorov identity** for some particular cases. Essentially, we are led to systems of differential equations for the $P_{jk}(t)$. In fact, we have already encountered some particular cases (see sections 2.1 and 2.3 although our concentration there was mainly on steady-state systems).

To introduce notations appropriate for the following sections, we choose an origin of time measurement and say that at epoch t > 0 the system is in state E_n if exactly n jumps occurred between 0 and t. Then, $P_n(t)$ equals the probability of the state E_n at epoch t, but $P_n(t)$ may be described also as the transition probability from an arbitrary state E_j at an arbitrary epoch s to the state E_{j+n} at epoch s+t.

2.7.2 The Poisson Process

In this case our postulates are:

- The process starts at epoch 0 from the state E_0 (meaning that $P_0(0) = 1$ and $P_n(0) = 0$ if $n > 0$))
- Direct transitions from a state E_j are possible only to state E_{j+1}
- Whatever the state Ej at epoch t, the probability of a jump within an ensuing short time interval between t and t+h equals λ + o(h), while the probability of more than one jump is o(h).

Note: We can interpret the states as the number of customers in a queue system, for example.

Assume $n \geq 1$ and consider the event that at epoch t+h the system is in state E_n. The probability of this event is $P_n(t+h)$ and the event can occur in three mutually exclusive ways:

(1) System in state E_n at epoch t and no jump occurs between t and t+h. Probability is $P_n(t)P_0(h) = Pn(t)[1-\lambda h] + o(h)$

(2) System is in E_{n-1} and exactly one jump occurs between t and t+h. Probability is $P_{n-1}(t)(\lambda h) + o(h)$

(3) Any other state at epoch t requires more than I jump between t and t+h. Probability is o(h).

Accordingly, we have

$$P_n(t+h) = P_n(t)[1-\lambda h] + P_{n-1}(t)(\lambda h) + o(h)$$

This may be re-written as

$$(P_n(t+h) - P_n(t))/h = -\lambda P_n(t) + \lambda P_{n-1}(t) + o(h)/h$$

As h -> 0, the last term tends to zero and we are left with the differential equation

$$d[P_n(t)]/dt = -\lambda P_n(t) + \lambda P_{n-1}(t)$$

If n = 0, possibilities (2) and (3) do not arise and we are let to

$$d[P_0(t)]/dt = -\lambda P_0(t)$$

This last equation with initial condition $P_0(0) = 1$ has solution $P_0(t) = e^{-\lambda t}$ (***check by substitution***)

It is an easy (enough) mathematical exercise to show that $P_1(t) = \lambda t e^{-\lambda t}$ using the initial condition $P_1(0) = 0$ and $P_0(t) = e^{-\lambda t}$. Continuing in this way one obtains the general expression for the Poisson distribution, $P_n(t) = (\lambda t)^n e^{-\lambda t}/n!$

2.7.3 The Pure Birth Process

This is the simplest generalization of the Poisson process and is obtained by permitting the probabilities of jumps to depend on the actual state of the system. This leads to the postulates:

- Direct transitions from a state E_j are possible only to state E_{j+1}
- If at epoch t the system is in state En, the probability of a jump within an ensuing short time interval between t and t+h equals λ_n + o(h), while the probability of more than one jump is o(h).

The same reasoning as in section 2.7.2 leads to the differential equation system

$$d[P_n(t)]/dt = -\lambda_n P_n(t) + \lambda_{n-1} P_{n-1}(t)$$

$$d[P_0(t)]/dt = -\lambda_0 P_0(t)$$

If it assumed that the system starts from an arbitrary initial state E_i then the initial conditions are

$$P_i(0) = 1,\ P_n(0) = 0 \text{ for all } n \neq i$$

According to (Feller, 1968) (page 449) explicit formulae for $P_n(t)$ have been derived by various authors. (Feller, 1968) also presents a particular case (called the Yule process) in which $\lambda_n = n\lambda$ which was proposed in connection with the mathematical theory of evolution. The rationale is that it is assumed that the probability that a member of a population gives birth (for example, by splitting) to a new member in a short time interval h is λh+o(h) so that if there are n members the probability that an increase takes place between t and t+h is nλh+o(h). In this case, the equations become

$$d[P_n(t)]/dt = -n\lambda P_n(t) + (n-1)\lambda P_{n-1}(t)$$

$$d[P_0(t)]/dt = 0$$

(Feller, 1968) presents the solution for this system, subject to the initial conditions given above.

2.7.4 The Birth-And-Death Process

This provides for transitions not only to higher but also to lower states.

This leads to the postulates:

- Direct transitions from a state E_n (n > 0) are possible only to its nearest neighbours, that is to state E_{n+1} or to state E_{n-1}. However, E_0 can only transit to E_1.
- If at epoch t the system is in state En, the probability of transit E_n -> E_{n+1} within an ensuing short time interval between t and t+h equals λ_n + o(h), while the probability

of transit E_n -> E_{n-1} within an ensuing short time interval between t and t+h equals μ_n + o(h).

- The probability of more than one change in the short time interval between t and t+h is o(h).

Following a similar approach as in 2.7.3 leads to the differential equation system,

$$d[P_n(t)]/dt = -(\lambda_n + \mu_n)P_n(t) + \lambda_{n-1}P_{n-1}(t) + \mu_{n+1}P_{n+1}(t) \qquad (n \geq 1)$$

$$d[P_0(t)]/dt = -\lambda_0 P_0(t) + \mu_1 P_1(t) \qquad (n = 0)$$

If the initial state is E_i then the initial conditions are

$$P_i(0) = 1, P_n(0) = 0 \text{ for all } n \neq i$$

Note: This is quite similar to the M/M/1 derivation in Section 2.3 although in that case we focussed more on the steady-state case. See also Section 2.5 where we had the steady-state form of the Birth-And-Death process.

Note on Steady-State: It can be shown that the limits

$$\lim_{t\to\infty} P_n(t) = P_n$$

exist and are independent of the initial conditions. The limiting (or Steady-State) probabilities P_n satisfy the equations

$$0 = -(\lambda_n + \mu_n)P_n + \lambda_{n-1}P_{n-1} + \mu_{n+1}P_{n+1} \qquad (n \geq 1)$$

$$0 = -\lambda_0 P_0 + \mu_1 P_1 \qquad (n = 0)$$

as previously given in section 2.5.

Note on correspondence to Markov Chains: This is discussed in (Feller, 1968) (page 457 etc). However, we will not go into this here except to point out the difference that changes can occur at arbitrary times whereas this is not true for the chain.

2.8 Problems

Question 1

At a tool service centre the arrival rate is two per hour and the service potential is three per hour. The hourly wage paid to the attendant at the service centre is €15 per hour and the hourly cost of a machinist away from his work is €40.

(i) Calculate the average number of machinists being served or waiting to be served at any given time

(ii) Calculate the average time a machinist spends waiting for service

(iii) Calculate the total cost of operating the system for an 8-hour day

(iv) Calculate the cost of the system if there were two attendants working together as a team (in effect acting together as a single server), each paid €15 per hour and each able to service on average two customers per hour.

Question 2

The manager of a bank must determine how many tellers, out of a maximum number of 6, should work on Fridays. For every minute a customer is in the bank, the manager believes that a cost of €0.05 is incurred. On average 2 customers per minute arrive at the bank, and a teller takes an average of 2 minutes to complete a customer's transaction. It costs the bank €9 per hour to hire a teller. Assuming an M/M/s queuing system, how many tellers should the bank have working on Fridays in order to minimise the sum of service costs and customer waiting costs?

3. Dynamic Programming (TBD)

This section is *TBD*.

Bibliography

Bolch, G., 2006. *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications.* 2 ed. s.l.:Wiley-Interscience.
Feller, W., 1968. *An Introduction to Probability Theory and Its Applications, Vol. 1.* 3 ed. s.l.:Wiley.
Hastie, T., Tibshirani, R. & Friedman, J., 2008. *The Elements of Statistical Learning.* Second ed. s.l.:Springer.
Jung, H. G., n.d. *Matrix Models,* s.l.: HanYang University.
Kleinrock, L., 1975. *Queueing Systems.* s.l.:Wiley.
McGarigal, K., n.d. *Models of landscape change? Landscape disturbance-succession models,* s.l.: s.n.
Montgomery, J., 2009. *Markov Chains,* s.l.: s.n.
Taha, H. A., 2007. *Operations research: an introduction.* 8 ed. s.l.:Pearson Prentice Hall.
Vega, M. E. F. d. l., 1998. *Approximate solutions for multi-server queueing systems with Erlangian service times and an application to air traffic management.* s.l.:MIT (Ph.D. thesis).