BLOCK I:

UNIT I - MARKOV CHAINS AND MARKOV PROCESSES

1.1 Introduction :

Dealing with uncertainty through random variable (discrete and continuous) is a tedious task for scientists since the 19^{th} century. For integral valued random variable s. It is often easy to apply powerful tools likegenerating functions end z- transforms. Instead of dealing with single random variable with correspondingprobability distribution(mass) function known, a family of random variables varying with time t(n-diorite)are dame in stochastic process

Thus stochastic process is a family of random variables, which varies with inspect to time (the parameter) and take specific values in a setor space.Real time space may be either discrete or continuous. In the ensuring section, We elucidate the concept of Markov Chain and its transition matrix. We also discuss the order of the MC. Higher order transition probabilities are also computed from Kolmogorov equation as well as transition probability matrix.

Definition:

Let { Xn,n ∈ T},Use stochastic process with state space E(discrete or continue) and time space T(discrete or continuous)

A thus a family of random variables { X t : t ∈ T} ,where T ={.....-1,0,1,2.....} or (-∞, ∞) or its subsets, takes its values from the state space E which is a subsets of real or complex space .

The collection of such processes consist of all kinds of stochastic processes that can be classified in to four different categories.

1) Discrete time, Discrete state space.
2) Discrete time, Continuous state space.
3) Continuous time, Discrete state space.
4) Continue time, Continuous state space

DS CS

$\{X_n : n \in N\}$ (1)	$\{X_n : n \in N\}$ (2)	DT
$\{X_t : t \in T\}$ (3)	$\{X_t : t \in T\}$ (4)	CT

Example1.

Let X_n denote the number of sixes up to the n^{th} throw of an unbiased die (6 faces) continuously. Then, clearly,$\{Xn : n \geq 0\}$ is a stochasticprocess with time space T= {0, 1, 2 …..}, and state space E = {0, 1, 2 ……}.

Example 2.

Consider the experiment of recording the temperature ata place at the end of every day. Let X_n denote the temperature measure on the n^{th} day, then $\{X_n : n \geq 0\}$ is a stochastic process with state space T = $(-\infty, \infty)$ (sometimes, the temp freezes below0^0) .

Example (3)

Let$X_t$denote the number of phone calls received at a telephone exchange board upto time t, That is the number of calls received during the interval [0, t), starting with initial time point t = 0, Then clearly $\{X_t : t \in T\}$ is a stochastic process with continuous time space$T = [0, \infty)$ and discrete state space E = $\{0, 1, 2 \dots\}$.

Example (4)

Consider the experiment of observing the price of gold in the whole sale market with initial time point t = 0. Let X_t denote the price of gold at time t (clock time). Then clearly $\{X_t : t \in T\}$ in a stochastic process with time space T=$(0, \infty)$ and stole spec E =$(0, \infty)$.

All the above examples are taken from real life situation and the classification of stochastic process is vivid from these examples.Next we see some of the processes with special properties.

1.2 Stochastic Processes – Independent increaments

Consider a stochastic process $\{X_t : t \in T\}$ with continuous time space T =$(-\infty, \infty)$. If for all t_1, $t_2 \dots . . t_n \in T t_1 < t_2 < \dots \dots < t_n$. The random variables,X (t_2) – X(t_1), X(t_3) $-$ X $(t_2$)…. , X(t_n) – X(t_{n-1}) are independent then,$\{X_t : t \in T\}$ is said to be a stochastic process with independent increments.

In the case of discrete parameter, $\{X_n : n \in N\}$ is a stochastic process, satisfying the Markovcondition.

Let Z_0 = X_0,Z_i = $X_i - X_{i-1}$,i = 1, 2…… , where N = { 0, 1, 2…..} be independent random variables, then $\{X_t : t \in T\}$ is the Stochastic Process with independent increments. Then the sequence of independent random variables$\{Z_n : n \geq 0\}$ is a stochastic process with independent increments.

Let $\{X_t : t \in T\}$ be a stochastic process with time space T = $(-\infty, \infty)$, and state space E = $(-\infty, \infty)$ (continuous time, continuous state space)

If for a given value,$X_{(s)}$, the value of $X_{(t)}$, t >s, do not depend as the values $X_{(u)}$, u<s, then the process $\{X_t : t \in T\}$ is said to be a Markov Process.

In mathematical form(probability distribution), this Markov Process can be defined as follows:

Iffort_1<t_2< ….. <t_n< t, $P_r\{\alpha \leq X_t \leq \beta \mid X_{t_1} = x_1, X_{t_2} =$ $x2,\ldots,Xtn = xn =$

$P_r\{\alpha \leq X_t \leq \beta \mid X_{t_n} = x_n\}$, then the process $\{X_t : t \in T\}$ is called Markov Process.

1.3. Markov Chains

Markov Chain: The discrete parameter Markov process $\{X_n : n \in N\}$ is known as Markov Chain with state space either discrete or continuous.

Consider a simple coin tossing experiment repeated for a number of times (costively), Twopossible outcomes for each trial are 'Head' and 'Tail'. Assume that Head occurs with probability p and that Tail occurs with probability q, so that p + q = 1.

Let us denote the outcomes of the n^{th} toss of the unbiased coin by X_n .

Then $X_n = \begin{cases} 1 & if\ \ head\ occurs \\ 0\ \ if\ tail\ occurs, & \text{for } n = 1,2,3,\ldots. \end{cases}$

That is

P_r { X_n = 1 } = p, and P_r $h\{X_n = 0\}$ = q. Hence the sequence of random variables, X_1,X_2……..,

Cen be written as $\{X_n : n \geq 1\}$, which is a Markov chain.

Definition:

The stochastic process $\{X_n : n = 0, 1, 2 \ldots.\}$ or $\{X_n : n \in N\}$, where No = { 0, 1, 2….}

is a Markov Chain if for i, j, $i_0, i_1, \ldots\ldots i_{n-1} \in N$,(or a subset of z).

$P_r = \{X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots\ldots X_1 = 1, X_0 = i_0\} = P_r\{X_{n+1} = j \mid X_n = i\} = P_{ij}$

whenever, the initial random variable X_0 is defined.

Here X_n =j means the outcome of the process in the n^{th} trial is j.

Remark:

1. The transition probability p_{ij} may or may not be independent of n. $\left(ie\, p_{ij} = p_{ij}{}^{(n)}.\right)$.

1.4.Transition Probabilities– Order of Markov Chains :

Consider a M. C. $\{X_n : n \geq 0\}$, then the m−step transition probability denoted $p_{ij}{}^{(m)}$ is defined as $P_{ij}{}^{(m)} = \{X_{n+m} = j \mid X_n = i\}$.

Transition Probability Matrix:

When m=1,the one step transition probabilities (p_{ij}), satisfies $p_{ij} \geq 0$ and $\sum_{j-0}^{\infty} p_{ij} = 1$ for all i = 0, 1,2,3,…

The transition probabilities for different state transitions may be written in*Matrix* from as follows:

$$P = \begin{pmatrix} P_{00} & P_{01} & P_{02} & -- \\ P_{10} & P_{11} & P_{12} & -- \\ - & - & - & -- \\ - & - & - & -- \end{pmatrix}$$

This matrix P is called a transition probability matrix (tpm) of the Markov Chain $\{X_n : n \geq 0\}$.

Example (5):

Consider a simple queuing system, before a counter designed for customer service. Customers arrive for service, to the counter (one server) who serves one customer at time epochs 0, 1, 2,…

Let Y_n denote the random variable, representing the number of customersarrive the counter during the time interval (n,n+1) for n = 0, 1, 2, …..

ClearlyY_n, are independent and identically distributed random variables, withprobability distribution P_r { Y_n = n } = p_k, k = 0, 1, 2, …… . Assume that the waiting room can accommodate only Mcustomers, including one in the counter.

Let $X_n$be the number of customers present at epoch n, including the one being served,if any, Then $\{X_n : n \geq 0\}$Is a Markov Chain with state space E = { 0, 1, 2, ……….,M }.

Now we have,

$$X_{n+1} = \begin{cases} Y_n \ \ if\ X_n = 0\ are\ 0 \leq \ Y_n \leq M \\ X_n + Y_n \ - 1\ if \quad 1 \leq X_n \ \leq M \\ \quad and\ 0 \leq Y_n \leq M + 1 - X_n \\ M \qquad\qquad\qquad otherwise. \end{cases}$$

The corresponding tpm is denoted by

$$P = \begin{bmatrix} & & q_0 q_1 q_3 & - & - & q_{M-2} q_{M-1} Q_M & & \\ & q_0 q_1 q_2 & & - & - & q_{M-2} q_{M-1} & Q_M & \\ & 0 & q_0 q_1 & - & - & q_{M-1} q_M Q_{M-1} & & \\ - & - & - & - & - & - & - & - \\ - & - & - & - & - & - & - & - \\ & 0 & 0 & - & - & - & q_0 q_1 Q_2 & \\ 0 & 0 & - & - & - & 0 & q_0 Q_1 & \end{bmatrix}$$

where, $Q_M = q_M + q_{M+1} + \ldots\ldots$ and $Q_o = 1$.

Example6.

Consider a particle moving back and front, in random fashion (random walk)along astraight line (lane) having absorbing barriers say 0 and 4. We describe its movement as follows:

When the particle is at position state) r (0< r < 4), it moves to state (r +1) with probability p or to state(r-1) with probability q, where p + q =1.

But as soon as the particle reached 0 or 4 it remains there itself (absorbing).

Example 7:

A particle performs a random walk with absorbing barriers, as 0 and 4. Whenever it is at any position r $(0 < r < 4)$, it moves to r + 1 with probability

P or to $(r-1)$ with probability q, p + q =1. But as soon as it reaches 0 or 4 it remains there itself. Let X_n be the position of the particle after n moves. The different states of X_n are different positions of the particle. $\{X_n\}$ is a Markov chain whose unit – step transition probabilities are given by

$$\Pr\{X_{n+1} = r + 1 \mid X_n = r\} = p$$

$$\Pr\{X_{n+1} = r - 1 \mid X_n = r\} = q \qquad 0 < r < 4$$

and

$$\Pr\{X_{n+1} = 0 \mid X_n = 0\} = 1,$$

$$\Pr\{X_{n+1} = 4 \mid X_n = 4\} = 1.$$

The transition matrix is given by States of X_{n+1}

$$\text{States of } X_n \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

1.4.1.General random walk between two barriers:

Consider that a particle that may be at any one position r, r = 0, 1,…., k (≥ 1) of the x – axis. From state r it moves to state r + 1, 1 ≤ r ≤ k – 1 with probability p and to state r – 1 with probability q. As soon as it reaches state 0 it remains there with probability a and is reflected to state 1 with probability 1 − a (0 < a < 1); if it reaches the state k it remains there with probability b and is reflected to k – 1 with probability 1 – b (0 < b < 1) . Then,$\{X_n\}$, where X_n is the position of the particle after n steps or moves, is Markov chain with state space S = { 0, 1,….k }. The transition matrix is

$$P = \begin{bmatrix} a & 1-a & 0 & - & 0 & 0 & 0 \\ q & 0 & - & - & 0 & 0 & 0 \\ - & - & - & - & - & - & - \\ 0 & 0 & 0 & - & q & 0 & p \\ 0 & 0 & 0 & - & 0 & 1-b & b \end{bmatrix}$$

If a = 1, then is an absorbing barrier and if a = 0, then is a reflecting barrier, if 0 < a < 1, 0 is an elastic barrier. Similar is the case with state k. The case when both 0 and k are absorbing barriers corresponds to the familiar

Gambler's ruin problem (with total capital between the two gamblers amounting to k).

Example 8:

Suppose that a coin with probability p for showing a *head* (success) is tossed indefinitely. Let X_n denote the outcome of the n^{th}trial, be k, where k (= 0, 1,…n) denote that there is a run of k successes, i. e. the length of the uninterrupted block of heads is k. Clearly { X_n, n ≥ 0 } constitutes a Markov Chain, with unit – step transition probabilities

$$P_{jk} = \Pr\{X_{n+1} = k \mid X_n = j\} = p, \quad k = j + 1$$

$$= q, \quad k = 0$$

$$= 0, \quad \text{otherwise.}$$

The transition matrix is given by

$$\text{States of } X_n \quad \begin{array}{c} \text{States of } X_{n+1} \\ \begin{array}{ccccccc} 0 & 1 & 2 & \dots & k & k+1 & \dots \end{array} \\ \begin{bmatrix} q & p & o & \dots & o & o & \dots \\ q & o & p & \dots & o & o & \dots \\ q & o & o & \dots & . & . & \dots \\ . & . & . & \dots & . & . & \dots \\ . & . & . & \dots & . & . & \dots \\ . & . & .. & \dots & . & . & \dots \\ q & o & o & \dots & o & p & .. \\ . & . & . & \dots & . & . & .. \\ . & . & . & \dots & . & . & \dots \end{bmatrix} \end{array}$$

Example 9:**Partial sum of independent *random variables***:

Consider a series coin tossing experiments, where the outcomes of n^{th} trial are denoted by 1 (for a head) and 0 (for a tail). Let X_n be the random variable denoting the outcome ofn^{th} trial and S_n = X_1 +….+ X_n be the n^{th} partial sum. The possible values of S_n are 0, 1,…,n, i. e. the states of S_n are r, r = 0, 1,…. n, $\{S_n, n \geq 0\}$ is a Markov chain with transition matrix as given below.

A Markov chain $\{X_n, n \geq 0\}$ with k states, where k is finite, is said to a finite Markov chain. The transition probability matrix P is, in this case, a square matrix with k - rows and k -columns. Examples7 and 8 deals with finite Markov chains.

Transition matrix

States of S_{n+1}

States of S_n $\begin{bmatrix} q & p & 0 & \dots & 0 & 0 & \dots \\ 0 & q & p & \dots & 0 & 0 & \dots \\ 0 & 0 & q & \dots & 0 & 0 & \dots \\ . & . & . & \dots & . & . & \dots \\ . & . & . & \dots & . & . & \dots \\ . & . & .. & .. & . & . & \dots \\ 0 & 0 & 0 & \dots & q & p & \dots \\ . & . & . & \dots & . & . & .. \\ . & . & . & \dots & . & . & \dots \end{bmatrix}$

The number of states could however be infinite. When the possible values of X_n form a denumerable set, then the Markov chain is said to be denumerably infinite or denumerable and the chain is said to have a countable state space. Examples 1 (d) ane of denumerable Markov chains.

1.5HIGHER TRANSITION PROBABILITES

1.5.1 Chapman – Kolmogorov equation:

We have so far considered unit – step or one – step transition probabilities, the probability of X_n given X_{n-1}, i. e. the probability of the outcome at the nth step or trial given the outcome at the previous step; p_{jk} gives the probability of unit – step transition from the state j at a trial to the state k at the next following trial. The m – step transition probability is denoted by

$$\Pr \{X_{m+n} = k \mid X_n = j\} = p_{jk}^{(m)};$$

$p_{jk}^{(m)}$ gives the probability that from the state j at nth trial, the k is reached at (m + n)th trial in m steps, i. e. the probability of transition from the state j to the state k in exactly m steps. The number n does not occur in the r. h. s. of the relation and the chain is homogeneous. The one – step transition probabilities $p_{jk}^{(1)}$ are denoted by p_{jk} for simplicity. Consider

$$p_{jk}^{(2)} = \Pr \{X_{n+2} = k \mid X_n = j\}.$$

The state k can be reached from the state j in two steps through some intermediate state r. Consider a fixed value of r; we have

$$\Pr\{X_{n+2} = k, X_{n+1} = r \mid X_n = j\}$$

$$= \Pr\{X_{n+2} = k, X_{n+1} = r \mid X_n = j\} \Pr\{X_{n+1} = r \mid Xn = j$$

$$= p_{rk}^{(1)} p_{jr}^{(1)} = p_{jr} p_{rk}.$$

Since these intermediate state r can assume values r = 1, 2, …., we have

$$p_{jk}^{(2)} = \Pr\{X_{n+2} = k \mid X_n = j\} = \sum_r Pr\{X_{n+2} = k, X_{n+1} = r \mid X_n = j\}$$

$$= \sum_r p_{jr}\, p_{jk}$$

(summing over for all intermediate states).

By induction, we have

$$p_{jk}^{(m+1)} = \Pr\{X_{n+m+1} = k \mid X_n = j\}$$

$$= \sum_r Pr\{X_{n+m+1} = k \mid X_{n+m} = r\} \Pr\{X_{n+m} = r \mid X_n = j\}$$

$$= \sum_r Pr\, p_{jr}^{(m)}\ .$$

Similarly, we get

$$p_{jk}^{(m+1)} = \sum_r p_{jr}\, p_{jr}^{(m)}\ .$$

In general, we have

$$p_{jk}^{(m+n)} = \sum_r p_{rk}^{(n)}\, p_{jr}^{(m)} = \sum_r p_{jr}^{(n)}\, p_{rk}^{(m)}\ .$$

This equation is a special case of Chapman – Kolmogorov equation, which is satisfied by the transition probabilities of a Markov chain.

From the above argument , we get

$$p_{jk}^{(m+n)} \geq p_{jr}^{(m)} p_{rk}^{(n)}\text{ , for any r.} \quad \square$$

1.5.2 Remark: We can put the results in terms of transition matrices as follows. Let P = (p_{jk}) denote the transition matrix of the unit – step transition and $P^{(m)}$ = $(p_{jk}^{(m)})$ denote the m-step transition matrix. For m = 2, we have the matrix $P^{(2)}$ whose elements are given by. It follow that the elements of

$P^{(2)}$ are the elements of the matrix obtained by multiplying the matrix P by itself, i. e.

$$P^{(2)} = P.P = P^2.$$

Similarly,

$$P^{(m+1)} = P^{(m)}.\,\mathrm{P} = \mathrm{P}\,.\,P^{(m)}$$

and $$P^{(m+n)} = P^{(m)}\,.\,P^{(n)} = P^{(n)}\,.\,P^{(m)}\,.$$

It should be noted that there exist non – Markov chain whose transition probabilities satisfy Chapman – Kolmogorov equation (example, see Feller I, p. 423, Parzen p. 203).

Example 2.

Consider the Markov chain of Example 1(g) . The two – step transition matrix is given by

$$\begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{5}{8} & \frac{5}{16} & \frac{1}{16} \\ \frac{5}{16} & \frac{1}{2} & \frac{3}{16} \\ \frac{3}{16} & \frac{9}{16} & \frac{1}{4} \end{bmatrix}$$

Hence $p_{01}^{(2)} = \Pr\{X_{n+2} = 1 \mid X_n = 0\} = \frac{5}{16}$ for $n \geq 0$.

Thus $\Pr\{X_2 = 1 \mid X_0 = 0\} = \frac{5}{16}$,

And $\Pr\{X_2 = 1\,, X_0 = 0\} = \Pr\{X_2 = 1 \mid X_0 = 0\}\ \Pr\{X_0 = 0\}$

$$= \left(\frac{5}{16}\right).\left(\frac{1}{3}\right) = \frac{5}{48}\,.$$

Example 3.

Two – state Markov chain . Suppose that the probability of a dry day (state 0) following a rainy day (state 1) is $\frac{1}{3}$ and that the probability of a rainy day following a dry day is $\frac{1}{2}$. We have a two – state Markov chain such that $p_{10} = \frac{1}{3}$ and $p_{01} = \frac{1}{2}$ and t. p. m.

$$\mathrm{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

We have $P^2 = \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \\ \frac{7}{18} & \frac{11}{18} \end{bmatrix} . p^4 = \begin{bmatrix} \frac{173}{432} & \frac{259}{432} \\ \frac{259}{648} & \frac{389}{648} \end{bmatrix}$

Given that 1 denote a dry day, the probability that May 3 is a dry day is $\frac{5}{12}$, and that May 5 is a dry is $\frac{173}{432}$ We can calculate the higher powers of P.

Example 4.

Consider a communication system which transmits the digits 0 and 1 through several stages. Let X_n, n ≥ 1 be the digit leaving the n^{th} stage of system and X_0 be the digit entering the first stage (leaving the 0^{th} stage). At each stage there is a constant probability q that the digit which enters will be transmitted unchanged (i. e. the digit will remain unchanged when it leaves), and probability p otherwise (i. e. the digit changes when it leaves), p +q = 1.

Here $\{X_n, n \geq 0\}$ is a homogeneous two – state Markov chain unit – step transition matrix

$$P = \begin{bmatrix} q & p \\ p & q \end{bmatrix}.$$

It can be shown (by mathematical induction or otherwise) that

$$p^m = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}(q-p)^m & \frac{1}{2} - \frac{1}{2}(q-p)^m \\ \frac{1}{2} - \frac{1}{2}(q-p)^m & \frac{1}{2} + \frac{1}{2}(q-p)^m \end{bmatrix}$$

Here $p_{00}^{(m)} = p_{11}^{(m)} = \frac{1}{2} + \frac{1}{2}(q-p)^m$

and $p_{01}^{(m)} = p_{10}^{(m)} = \frac{1}{2} - \frac{1}{2}(q-p)^m$.

Also as m → ∞, $\lim p_{00}^{(m)} = \lim p_{01}^{(m)} = \lim p_{10}^{(m)} = \lim p_{11}^{(m)} \to \frac{1}{2}$.

Suppose that the initial distribution is given by

Pr $\{X_0 = 0\}$ and Pr $\{X_0 = 1\}$ = b = 1 – a.

Then we have

Pr $\{X_m = 0, X_0 = 0\}$ = Pr $\{X_m = 0 \mid X_0 = 1\}$ Pr $\{X_0 = 0\}$

$= ap_{00}^{(m)}$

and Pr$\{X_m = 0, X_0 = 1\} = bp_{10}^{(m)}$.

The probability that the digit entering the first stage is 0 given that the digit leaving the mth stage is 0 can evaluated by applying Bayes' rule. We have

$$\Pr \quad \{X_0 = 0 \mid X_m = 1\}$$

$$= \frac{\Pr\{X_m=0 \mid X_0=0\}\Pr\{X_0=0\}}{\Pr\{X_m=0 \mid X_0=0\}\Pr\{X_0=0\}+\Pr\{X_m=0 \mid X_0=1\}\Pr\{X_0=1\}}$$

$$= \frac{ap_{(00)}^{(m)}}{ap_{(00)}^{(m)}+bp_{(01)}^{(m)}}$$

$$= \frac{a\left\{\frac{1}{2}+\frac{1}{2}(q-p)^m\right\}}{a\left\{\frac{1}{2}+\frac{1}{2}(q-p)^m\right\}+b\left\{\frac{1}{2}-\frac{1}{2}(q-p)^m\right\}}$$

$$= \frac{a\{1+(q-p)^m\}}{1+(a-b)(q-b)^m}.$$

1.5.3 Probability distribution- definition

Probability distribution of random variables involved in a markov chain can be studied in this section.The joint distribution of consecutive random variables can be found using the following techniques:
It may be seen that the probability distribution of Random variables X_r, X_{r+1},……, X_{r+n} can be computed in terms of the transition probabilities p_{jk} and the initial distribution of X_r, is known .Suppose, for simplicity, take r = 0, then

$$\Pr\{X_0 = a,\ X_1 = b, \ldots,\ X_{n-1} = j,\ X_n = k\}$$

$$= \Pr\{X_n = k \mid X_{n-1} = j, \ldots,\ X_o = a\} \Pr \quad \{X_{n-1} = j, \ldots,\ X_o = a\}$$

$$= \Pr\ \{X_n = k \mid X_{n-1} = j\}\ \Pr\ \{X_n = j \mid X_{n-2} = i\}\ \Pr\ \{X_{n-2} = i, \ldots, Xo=a$$

$$= \Pr\ \{X_n = k \mid X_{n-1} = j\} \quad \Pr\ \{X_n = j \mid X_{n-2} = i\}\ \ldots\ \Pr\ \{X_1 = b \mid X0=a \Pr X0=a$$

$$= \{\ Pr(X_0 = a)\}p_{ab} \ldots p_{ij}p_{jk},$$

Thus,

$$\Pr\{X_r = a,\ X_{r+1} = b, \ldots,\ X_{r+n-2} = i,\ X_{r+n-1} = j, X_{r+n=k}\}$$

$$= \{\ Pr\ (X_r = a)\}p_{ab} \ldots p_{ij}p_{jk},$$

Example 1:

Let $\{X_n, n \geq 0\}$ be a Markov chain with three states 0, 1, 2 and with transition matrix

$$\begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix},$$ and the initial distribution Pr $\{X_0 = i\} = \frac{1}{3}$, I = 0, 1, 2.

we have Pr $\{X_1 = 1 \mid X_0 = 2\} = \frac{3}{4}$

Pr $\{X_2 = 2 \mid X_1 = 1\} = \frac{1}{4}$

Pr $\{X_2 = 2 \mid X_1 = 1 \mid X_0 = 2\}$

$= \Pr \{X_2 = 2 \mid X_1 = 1\} \Pr \{X_1 = 1 \mid X_0 = 2\} = \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16}$

Pr $\{X_2 = 2, X_1 = 1, X_0 = 2\}$

$= \Pr \{X_2 = 2, X_1 = 1 \mid X_0 = 2\} \Pr \{X_0 = 2\} = \frac{3}{16} \cdot \frac{1}{3} = \frac{1}{16}$

Pr $\{X_3 = 1, X_2 = 2, X_1 = 1, X_0 = 2\}$

$= \Pr \{X_3 = 1 \mid X_2 = 2, X_1 = 1, X_0 = 2\} \times \Pr \{X_2 = 2, X_1 = 1, X_0 = 2\}$

$= \Pr \{X_3 = 1 \mid X_2 = 2\} \left(\frac{1}{16}\right) = \frac{3}{4} \cdot \frac{1}{16} = \frac{3}{64}.$

1.5.4.Remark

The matrix of transition probabilities together with initial distribution , completely specifies a Markov Chain $\{X_n : n = 0, 1, 2, \ldots\}$.

We state (without proof) the general existence theorem of Markov chains.

Given the state space E and the sequence of stochastic matrices ($p_{jk}^{(n)}$)= $P^{(n)}$,there exist a MarkovChain $\{X_n, n \geq 0\}$ with state space E and transition probability matrix, $P^{(n)}$. (For proof, see Iosifescu & Tautu, Stochastic Processes – I, Springer – Verlag (1973), Chung (1967)).

UNIT II

CLASSIFICATION OF STATES OF STATES AND CHAINS

2.1 Classification of States

2.2. Markov Chains with denumerable number of states

2.3 Reducible Chains

2.1. Classification of states:

The states j, j = 0, 1, 2, …. Of a Markov chain $\{X_n, n \geq 0\}$ can often be classified in a distinctive manner according to some fundamental properties of the system. By means of such classification it is possible to identify certain types of chains.

Communication Relations

If $p_{ij}^{(n)} > 0$ for some n ≥ 1, then we say that state j can be reached or state j is accessible from state I; the relation is denoted by 1 → j . Conversely, if for all n, $p_{ij}^{(n)}$ = 0, then j is not accessible from I; in notation I ↛ j.

If two states I and j are such that each is accessible from the other then we say that the two states communicate; it is denoted by I ⟷ j ; then there exist integer m and n such that

$$p_{ij}^{(n)} > 0 \text{ and } p_{ij}^{(m)} > 0 .$$

The relation → is transitive, i. e. if i → j and j → k then i → k. From Chapman – Kolmogorov equation

$$p_{ik}^{(m+n)} = \sum_r p_{ir}^{(m)} \, p_{rk}^{(n)}$$

$$p_{ik}^{(m+n)} \geq p_{ij}^{(m)} p_{jk}^{(n)}$$

where the transitivity property follows.

The relation ⟷ is also transitive; i. e. i ⟷ j, j ⟷ k imply I ⟷ k.

The relation is clearly symmetric, i. e. if i ⟷ j, then j ⟷ i.

The digraph of a chain helps in studying the communication relations.

From we see that $0 \leftrightarrow 1$ and $1 \leftrightarrow 2$ implies $0 \leftrightarrow 2$.

The states of this chain are such that every state can be reached from every other state.

2. 1. 2 Class Property

A class of states is a subset of the state space such that every of the class communicates with every other and there is no other state outside the class which communicates with all other states in the class. A property defined for all states of a chain is a class property is its possession by one state in a class implies its possession by all states of the same class. One such property is the periodicity of a state. Periodicity:

State I is a return state if $p_{ii}^{(n)} \geq 0$ for some n ≥ 1. The period d_i of a return to state i is defined as the greatest common divisor of all m such that $p_{ii}^{(m)} > 0$. Thus

$$d_i = \text{G. C. D.} \left\{ m : p_{ii}^{(m)} > 0 \right\} ;$$

State i is said to be aperiodic if $d_i = 1$ and periodic if $d_i > 1$. Clearly state I is aperiodic if $p_{ii} \neq 0$.

It can be shown that two distinctive states belonging to the same class have same period.

2. 1. 3 Classification of Chains:

If C is a set states such that no state outside C can be reached from any state in C, then C is said to be closed. If C is closed and j ∈ C while k ∉ C, then $p_{jk}^{(n)} = 0$ for all, i. e. C is closed iff $\sum_{j \in c} p_{ij} = 1$ for every I ∈ C. Then the sub-matrix $P_1 = (p_{ij})$, I, j, ∈ C, is also stochastic and P can be expressed in the canonical form as :

$$P = \begin{bmatrix} P_1 & 0 \\ R_1 & Q \end{bmatrix}$$

A closed set may contain one or more states. If a closed set contains only one state j then state j is said to be absorbing: j is absorbing iff $p_{jj} = 1$, $p_{jk} = 0$, k ≠ j. In Example 1 (b), states 0 and 4 are absorbing.

Every finite Markov chain contains at least one closed set, i. e. the set of all states or the state space. If the chain does not contain any other proper closed subset other than the state space, then chain is called irreducible; the t. p. m. of irreducible chain is an irreducible matrix. In an irreducible Markov chain every state can be reached from every other state. The Markov chain of Example1(g) is irreducible. Chains which are irreducible are said to be reducible or non – irreducible; the t. p. m. is reducible. The irreducible matrices may be subdivided into two classes: primitive (aperiodic) and unprimitive (cyclic or periodic) (See Section A.4 Appendix). A Markov chain is primitive (aperiodic) iff the corresponding t. p. m. is primitive. In an irreducible chain states belong to the same class.

2. 1. 4 **Transient and Recurrent States**

We now proceed to obtain a more sensitive classification of the states of a Markov chain.

Suppose that a system starts with state j. Let $f_{jk}^{(n)}$ be the probability that it reaches the state k for the first time at the nth step (or after n transitions) and let $p_{jk}^{(n)}$ be the probability that it reaches state k (not necessarily for the first time) after n transitions. Let τ_k given that the chain starts at state j. A relation can be established between $f_{jk}^{(n)}$ and $p_{jk}^{(n)}$ as follows. The relation allows $f_{jk}^{(n)}$ to be expressed in terms of $p_{jk}^{(n)}$.

Theorem2.1.5(First Entrance Theorem)

Whatever be the states j and k, $p_{jk}^{(n)} = \sum_{r=0}^{n} f_{jk}^{(r)}\, p_{kk}^{(n-r)}$, n ≥ 1, with $p_{kk}^{(0)} = 1$, $f_{jk}^{(0)} = 0$, $f_{jk}^{(1)} = p_{jk}$.

Proof: Intuitively, the probability that starting, with j, state k is reached for the first time at the r^{th} step and again after that are $(n-r)^{th}$ step is given by $f_{jk}^{(n)} p_{kk}^{(n-r)}$ for all r ≤ n. These cases are mutually exclusive . Hence the result

Note: (1) The recursive relation (4.1) can also be written as

$$p_{jk}^{(n)} = \sum_{r=1}^{n-1} f_{jk}^{(r)}\, p_{kk}^{(n-r)} + f_{jk}^{(n)}, \quad n > 1.$$

(2) For a rigorous proof which uses the strong Markov property, see Iosifescu (1980).

(3) In practice, it is sometimes convenient to compute $f_{jk}^{(r)}$ from the diagraph of chain.

2. 1. 6. First passage time distribution

Let F_{jk} denote the probability that starting with state j the system will ever reach state k. Clearly

$$F_{jk} = \sum_{n=1}^{\infty} f_{jk}^{(n)}$$

We have $$\sup p_{jk}^{(n)} \leq F_{jk} \leq \sum_{m \geq 1} p_{jk}^{(m)} \text{ for all } n \geq 1.$$

We have to consider two cases, F_{jk} = 1 and F_{jk}< 1.

When F_{jk} = 1, it is certain that the system starting with state j will reach state k; in this case { $f_{jk}^{(n)}$, n = 1, 2, … } is a proper probability distribution and this first passing time distribution for k given that the system starts with j.

The mean (first passing) time from state j to state k is given by $\mu_{jk} = \sum_{n=1}^{\infty} n f_{jk}^{(n)}$.

In particular, when k = j, { $f_{jj}^{(n)}$, n = 1, 2,…} represents the distribution of the recurrence times of j; and F_{jj} = 1 will imply that the return to the state j is certain . In this case

$\mu_{jj} = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$ is known as the mean recurrence time for the state j.

Thus, two questions arise concerning state j: first, whether the return to state j is certain and secondly, when this happens, whether the mean recurrence time μ_{jj} is finite.

It can be shown that

$$d_i = \text{G.C.D. } \left\{m : p_{ii}^{(m)} > 0\right\} = \text{G. C. D. } \left\{m : f_{ii}^{(m)} > 0\right.$$

Definitions

A state j is said to persistent (the word recurrent is also used by some authors; we shall however use the word persistent) if F_{jk}< 1 (i. e. return to state j is uncertain). A persistent state j is said to be null persistent if μ_{jj} = ∞, i. e. if the mean recurrence time is infinite and is said to be non – null (or positive) persistent if μ_{jj}<∞ ,

Thus the states of Markov chain can be classified as transient and persistent, and persistent states can be subdivided as non – null and null persistent.

A persistent non – null and aperiodic state of a Markov chain is said to be ergodic. Consider the following example.

Example 5.

Let $\{X_n, n \geq 0\}$ be a Markov chain having state S = { 1, 2, 3, 4) and transition matrix

$$P = \begin{bmatrix} \frac{1}{3} \frac{2}{3} & 0 & 0 \\ 1 \quad 0 & 0 & 0 \\ \frac{1}{2} \quad 0 & \frac{1}{2} & 0 \\ 0 \quad 0 & \frac{1}{2} \frac{1}{2} \end{bmatrix}$$

Here $f_{33}^{(1)} = \frac{1}{2}$, $f_{33}^{(2)} = f_{33}^{(3)} = ... = 0$ so that $F_{33} = \sum_{n=1}^{\infty} f_{33}{}^{(n)} = \frac{1}{2} + 0 + 0..=1/2 < 1$.

Hence state 3 is transient.

Again $f_{44}^{(1)} = \frac{1}{2}$, $f_{\ 44}^{(n)} = 0$, n ≥ 2, so that $F_{44} = \sum_{n=1}^{\infty} f_{44}{}^{(n)} = \frac{1}{2} + 0 + 0 ... = 1/2 < 1$.

Hence state 4 is also transient.

For state 1:

Now $f_{11}^{(1)} = \frac{1}{3}$, $f_{11}^{(2)} = \frac{2}{3}$ and $F_{11} = \sum_{n=1}^{\infty} f_{11}{}^{(n)} \frac{1}{3} + \frac{2}{3} = 1$, so that state 1 is persistent.

Further since $\mu_{11} = = \sum_{n=1}^{\infty} n f_{11}^{(n)}\ 1.\ \frac{1}{3} + 2.\ \frac{2}{3} = \frac{5}{3}$, state 1 is non – null persistent.

Again $p_{11} = \frac{1}{3} > 0$, so that state 1 is aperiodic. Since state 1 is non-null persistent and aperiodic clearly State 1 is *ergodic.*

For State 2:

$$f_{22}^{(1)} = 0, f_{22}^{(2)} = 1 \cdot \frac{2}{3}, f_{22}^{(3)} = 1 \cdot \frac{1}{3} \cdot \frac{2}{3}, f_{22}^{(4)} = 1 \cdot \left(\frac{1}{3}\right)^2 \cdot \frac{2}{3}$$

$$\ldots f_{22}^{(n)} = 1 \cdot \left(\frac{1}{3}\right)^{n-2} \cdot \frac{2}{3}, \text{ n} \geq 2$$

So that $F_{22} = \sum_{n=1}^{\infty} f_{22}^{(n)} = \sum_{k=2}^{\infty} \left(\frac{1}{3}\right)^{k-2} . \frac{2}{3} = 1$

Thus state 2 is persistent.

Now we have

$$\mu_{22} = \sum_{k=1}^{\infty} k\, f_{22}^{(k)} = \sum_{k=2}^{\infty} k \left(\frac{1}{3}\right)^{k-2} . \frac{2}{3} = 2\sum_{k=2}^{\infty} k \left(\frac{1}{3}\right)^{k-1} = \frac{5}{2}$$

So that state 2 is non – null persistent. It is also aperiodic, and hence *ergodic.*

In the above example , calculation of $f_{ii}^{(n)}$ and so of $F_{ii} = \sum f_{ii}^{(n)}$ was easy. But sometimes it is not so easy to calculate $f_{ii}^{(n)}$ for n ≥ 2. In view of this, another characterization of persistence is given in Theorem 2. 2.

Example 6:

Consider a Markov chain with transition matrix

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{2} \end{pmatrix}.$$

Show that all states of the above MC are ergodic.

Solution: It can be easily seen that the chain is irreducible. Consider state 4: we have $p_{44} = \frac{1}{2} > 0$; state is aperiodic and $f_{44}^{(1)} = \frac{1}{2}$, $f_{44}^{(2)} = \frac{1}{8}$, $f_{44}^{(3)} = \frac{1}{8}$, $f_{44}^{(4)} = \frac{1}{4}$, $f_{44}^{(n)} = 0$, n > 4 so that $F_{44} = 1$ and $\mu_{44} = 1. \frac{1}{2} + 2. \frac{1}{8} + 3. \frac{1}{8} + 4. \frac{1}{4} = \frac{17}{8} < \infty$. Thus state 4 is ergodic. Hence all states are ergodic.

Theorem 2.1.7

Consider a M. C, $\{ X_n : n \geq 0 \}$. Then the state j is persistent iff

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty . \qquad (4.6)$$

Proof: Let P_{jj} (s) = $\sum_{n=0}^{\infty} p_{jj}^{(n)} s^n$ = 1 + $\sum_{n=1}^{\infty} p_{jj}^{(n)} s^n$, |s| < 1

and F_{jj} (s) = $\sum_{n=0}^{\infty} f_{jj}^{(n)} s^n$ = $\sum_{n=1}^{\infty} f_{jj}^{(n)} s^n$, |s| < 1

be the generating functions of the sequences $\left\{p_{jj}^{(n)}\right\}$ and $\left\{f_{jj}^{(n)}\right\}$ respectively.

We have from(4.1)

$$p_{jj}^{(n)} = \sum_{r=0}^{\infty} f_{jj}^{(r)} p_{jj}^{(n-r)} \quad (4.\ 7)$$

Multiplying both sides of (4. 7) by s^n and adding for all n ≥ 1, we get

$$P_{jj}\ (s) - 1 = F_{jj}\ (s)\ P_{jj}\ (s).$$

The right hand side of the above is immediately obtained by considering the fact that the r. h. s of (4. 7) is a convolution of { F_{jj}} and { P_{jj} } and that the generating of the convolution is the product of the two generating functions. Thus we have

$$P_{jj}\ (s) = \frac{1}{1- F_{jj}\ (s)},\ |s| < 1.$$

Assume that state j is persistent which implies that F_{jj} = 1. Using Abel's lemma, we get

$$lim_{s\to 1} F_{jj}(s) = 1$$

Thus $lim_{s\to 1} F_{jj}(s) \to \infty$

Since the coefficients of P_{jj} (s) are non – negative Abel's lemma applies and we get $\sum p_{jj}^{(n)} = \infty$, Conversely, suppose the state j is transient, then by Abel's lemma, we get

$$lim_{s\to 1} F_{jj}(s) < 1$$

Also from (4.8), $lim_{s\to 1} P_{jj}(s) < \infty$

Since the coefficients $P_{jj}^{(n)} \geq 0$, we get

$$\sum_n P_{jj}^{(n)} < \infty.$$

This is a contradiction to our hypothesis. ↕. Hence the j is persistent. □

2.1.8.Remark:

The result of Theorem 2. 2 can also deduced from the following result. Doblin's Formula: Whatever be the states j and k,

$$F_{jk} = lim_{m\to\infty} \frac{\sum_{n=1}^{m} P_{jk}^{(n)}}{1+\sum_{n=1}^{m} P_{kk}^{(n)}}$$

(4. 9)

And, in particular F_{jj} = 1 - $lim_{m\to\infty} \frac{1}{1+\sum_{n=1}^{m} P_{jj}^{(n)}}$.

Example 7:

Consider the Markov chain with t. p. m.

$$P = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$$

The chain is irreducible as the matrix is so. We have

$$P^2 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} . P^3 = P ;$$

In general, $P^{2n} = P^2$, P^{2n+1} = P ,

So that $p_{ii}^{(2n)} > 0$, $p_{ii}^{(2n+1)}$ = 0 for each I .

The states are periodic with period 2.

We find that f_{11} = o, $f_{11}^{(2)}$ = 1 so that $f_{11} = \sum_n f_{11}^{(n)}$ = 1, i. e. state 1 is persistent and hence the other states 0 and 2 are also persistent.

Now μ_{11} = = $\sum_n f_{11}^{(n)}$ = 2,

i. e. state is non - null. Thus the states of the chain are periodic (each with period 2) and persistent non – null. Further,

$$p_{11}^{(2n)} \to \frac{t}{\mu_{11}} = \frac{2}{2} = 1 \text{ for all n.}$$

We now state a lemma without proof (for proof, see Feller, Vol. I).

2.1.9.Basic limit theorem of renewal theory:

Lemma 2. 1. Let $\{f_n\}$ be a sequence such that $f_n \geq 0$, $\sum f_n = 1$ and t ($\geq$ 1) be the greatest common divisor of those n for which $f_n > 0$.

Let $\{u_n\}$ be another sequence such that $u_0 = 1$ and $u_n = \sum_{r=1}^{n} f_r u_{n-r} (n \geq 1)$. Then

$$lim_{n \to \infty} u_{nt} = \frac{t}{\mu}, \qquad (4.\ 10)$$

Where $\mu = \sum_{n=1}^{\infty} n f_n$, the limit being zero when $\mu = \infty$; and $lim_{N \to \infty} u_N = 0$ whenever N is not divisible by t. The lemma will be used to prove some important results.

Theorem 2.1.10:

If state j is persistent non – null, then as n → ∞

(i)$p_{jj}^{(nt)} \to \frac{t}{\mu_{jj}}$,when state j is periodic with period r;

and (ii) $p_{jj}^{(n)} \to \frac{1}{\mu_{jj}}$,when state j is aperiodic.

In case state j is persistent null, (whether periodic or aperiodic), then

$$p_{jj}^{(n)} \to 0 \text{ , as n} \to \infty.$$

Proof: Let state j be persistent; then

$$\mu_{jj} = \sum_n n\, f_{jj}^{(n)} \quad \text{is defined.}$$

Since (4.7) holds, we may put

$$f_{jj}^{(n)} \text{ for } f_n \text{ , } p_{jj}^{(n)} \text{ for } u_n, \text{ and } \mu_{jj} \text{ for } \mu$$

Applying the lemma(2.1), we get

$p_{jj}^{(nt)} \to \frac{t}{\mu_{jj}}$, as n $\to \infty$, when state j is periodic with period t.

When state j is aperiodic (i. e. t =1), $p_{jj}^{(n)} \to \frac{1}{\mu_{jj}}$, as n $\to \infty$.

In case state j is persistent null, $\mu_{jj} = \infty$, and $p_{jj}^{(n)} \to 0$ as n $\to \infty$.

2.1.11Note:

(1) If j is persistent non – null, then $lim_{n\to\infty} p_{jj}^{(n)} > 0$

and (2) if j is persistent null or transient then $lim_{n\to\infty} p_{jj}^{(n)} \to 0$.

Theorem 2.1.12

If state k is persistent null, then for every j, $lim_{n\to\infty} p_{jk}^{(n)} \to 0$. (4. 14)

If state k is aperiodic, persistent non – null then, $lim_{n\to\infty} p_{jk}^{(n)} \to \frac{F_{jk}}{\mu_{kk}}$. (4. 15)

Proof: We have

$$p_{jk}^{(n)} = \sum_{r=1}^{n} f_{jk}^{(r)} p_{kk}^{(n-r)} .$$

Let n > m, then

$$p_{jk}^{(n)} = \sum_{r=1}^{m} f_{jk}^{(r)} p_{kk}^{(n-r)} + = \sum_{r=m+1}^{n} f_{jk}^{(r)} p_{kk}^{(n-r)}$$

$$\leq \sum_{r=1}^{m} f_{jk}^{(r)} p_{kk}^{(n-r)} + \sum_{r=m+1}^{n} f_{jk}^{(r)} . \qquad (4.\ 16)$$

Since state k is persistent null,

$$p_{kk}^{(n-r)} \to 0, \text{ as } n \to \infty.$$

Further, since

$$\sum_{m=1}^{\infty} f_{jk}^{(m)} < \infty, \sum_{r=m+1}^{n} f_{jk}^{(r)} \to 0 \text{ as } n, m \to \infty.$$

Hence as n $\to \infty$

$$p_{jk}^{(n)} \to 0.$$

From (4. 16)

$$p_{jk}^{(n)} - \sum_{r=1}^{m} f_{jk}^{(r)} p_{kk}^{(n-r)} \leq \sum_{r=m+1}^{n} f_{jk}^{(r)}$$

(4. 16 a)

Since j is aperiodic, persistent and non – null, then by Theorem 2. 3.

$$p_{kk}^{(n-r)} \to \frac{1}{\mu_{kk}} \text{ as n} \to \infty.$$

Hence from (4.16 a), we get, as n, m → ∞, $p_{jk}^{(n)} \to \frac{F_{jk}}{\mu_{kk}}$ ☐.

2.2 MARKOV CHAIN WITH DENUMERABLE NUMBER OF STATES

So far we discussed Markov chains with finite number of states. The result can be generalized to chain with a denumerable number of states (or with countable state space). Let P = (p_{ij}) be the t. p. m. of the chain $\{X_n, n \geq 1$ with countable states space S = { 0, 1, 2,… }. Then *Pk = pij(k)* is well defined. The states of the chain may not constitute even single closed set. For example when

$$p_{ij} = 1, \text{ j = I + 1}$$

$$= 0, \text{ otherwise,}$$

The states do not belong to any closed set, including S.

For dealing with a chain with a countable state space, we need a more sensitive classification of states – transient, persistent null and persistent non – null. Beside irreducibility and aperiodicity, non – null persistence is required for ergodicity for such a chain (a chain with countable state space) while aperiodicity and irreducibility (or some type of reducibility) were enough for ergodicity for a finite chain. We shall state the theorem without proof.

Theorem 2.2.1 (General Erogodic Theorem)

Let $\{X_n\}$ be an irreducible, aperiodic Markov chain with state space S = {…., k,…} and having t.p.m. P =(p_{ij}). If the chain is transient or persistent non – null then $lim_{n\to\infty} p_{jk}^{(n)}$ = 0. If the chain is persistent non – null then limits $lim_{n\to\infty} p_{jk}^{(n)} = v_k$ exist and are independent vector. The vector V' is given by the solution of

$$V' = V' \text{ P,}$$

That is $\quad v_k = \sum_j v_j \, p_{jk}$.

Proof:

The above theorem is a generalization of Theorem 2. 7 concerning ergodicity of finite, irreducible, aperiodic chain

Example 8.

A Markov chain occurring in queueing theory is a chain with a countable state space S = { 0, 1, 2,…} and transition probability matrix

$$P = \begin{bmatrix} p_0 p_1 p_2 p_3 & - \\ p_0 p_1 p_2 p_3 & - \\ 0 \quad p_0 p_1 p_2 & - \\ 0 \quad p_0 p_1 p_2 & - \\ 0 \quad p_0 p_1 p_2 & - \end{bmatrix}$$

where $\sum p_k = 1$.

Let $P(s) = \sum_k p_k s^k$ and $V(s) = \sum_k v_k s^k$

Be generating functions of $\{p_k\}$ and $\{v_k\}$ respectively. Assume that $p_i > 0$ for all i. The chain is irreducible and aperiodic. It can be shown that the states are transient, persistent null or persistent non – null according as $P'(1) > 1, = 1$ or < 1. Thus when $P'(1) < 1$, v_k's are the unique solutions of the equations (6. 2). The equations (6. 2) becomes:

$$v_0 = p_0 v_0 + p_0 v_1$$

$$v_1 = p_1 v_0 + p_1 v_1 + p_0 v_2$$

$$v_2 = p_2 v_0 + p_2 v_1 + p_1 v_2 + p_0 v_3$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots. \quad \ldots$$

$$v_k = p_k v_0 + p_k v_1 + p_{k-1} v_2 + \ldots. + p_0 v_{k+1}$$

Multiplying both sides of the (k + 1) st equation by s^k (k = 0, 1, 2,…) and adding over k, we get

$$V(s) = v_0\, P(s) + v_1\, P(s) + v_2\, s\, P(s) + v_3 s^2\, P(s) + \ldots.$$

$$= P(s)\left\{v_0 + (V(s) - v_0) - s\right\}$$

This gives

$$V(s) = v_0(1 - s)P(s) \,/\, (P(s) - s),$$

In terms of v_0 which can be evaluated from $\sum v_k = 1$. We have

$$lim_{s\to 1}\frac{V(s)}{v_0} = \qquad lim_{s\to 1}\frac{\{1-s\}P(s)}{P(s)-s}$$

Whence $v_0 = 1 - P'(1)\ (> 0).$

Thus $V(s) = \frac{\{1-P'(1)\}(1-s)\,P(s)}{P(s)-s}$.

2.3 REDUCIBLE CHAINS

In the previous section, we studied the limiting properties of irreducible Markov Chains. In this section we propose to discuss some properties of reducible chains.

2.3.1 Finite Reducible chains with one closed set>

Consider a Markov chain with state space S , having a single closed set C ,in which all states communicate with each other. Also assume that the states of C are periodic. The ergodicity of finite irreducible Markov chainswas already considered in Theorem 2.7. Further ergodicity theorem for reducible chains having a single closed class of periodic states is given below.

Theorem 2.3.2 (Ergodic theorem for reducible chain)

Let $\{X_{:n} : n \geq 0\}$ be a finite Markov Chain with periodic states. Let P be the transition matrix of the m – state chain with state space S, and P_1 the transition (submatrix) of transitions among the k $(\leq m)$ members of the closed class C. Let V_1' = { …., $v_{j,....}$} be the stationary distribution corresponding to the stochastic submatrix P_1, i. e. $P_1^n \to e\, V_1'$ .If $V' = (V_1', 0')$, then, as $P^n \to$ e V_1'. In other words, elementwise V' is the stationary distribution corresponding to the matrix P.

Proof: An outline of proof is given below:

The transition matrix of the chain can be put in canonical from

$$P = \begin{bmatrix} P_1 & 0 \\ R_1 & Q \end{bmatrix}$$

Where the stochastic (sub) matrix corresponds to transitions among the members of class C and Q corresponds to transitions among the other the states (of S - C).

We have $P^n = \begin{bmatrix} P_1^n & 0 \\ R_n Q^n & \end{bmatrix}$

Where $R_n = R_{n-1}P_1 + Q^{n-1}R_1$. Writing R_1 = R, We get

$$R_{n+1} = \sum_{i=0}^{n} Q^i R P_1^{n-i} = \sum_{i=0}^{n} Q^{n-i} R P_1^i$$

As n → ∞ $\qquad P_1^n \to e\,V_1'$

and $\qquad Q^n \to 0.$

Again it can be shown that, as n → ∞

$$R_{n+1} \to e\,V_1'$$

So that, writing $V' = (V_1', 0')$ we have

$$P^n \to e\,V_1'.$$

2. 3. 3 Chain with one Single Class of Persistent Non – null Aperiodic States

Now suppose that the states of the closed class C are non – null persistent and aperiodic, the remaining states of S being transient; the transient states constitute a set T.

Then we have , for each pair i, j,

$$lim_{n\to\infty} p_{ij}^{(n)} = v_j$$

Is independent of i, when i, j are persistent, and also when j is persistent and i is transient; again

$$lim_{n\to\infty} p_{ij}^{(n)} = 0 \text{ when j is transient.}$$

In this case shall write the transition matrix as

$$P = \begin{bmatrix} P_1 & 0 \\ R_1 & M \end{bmatrix},$$

where M gives the matrix of transitions among the transient states.

Example 9: Consider a reducible chain with S = $\{1, 2, 3, 4\}$ and t. p. m.

$$P = \begin{bmatrix} P_1 & 0 \\ R_1 & Q \end{bmatrix}$$

Where $\qquad P_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}. \quad R_1 = \begin{bmatrix} 0 & \frac{1}{4} \\ 0 & \frac{1}{4} \end{bmatrix}. \quad M = \begin{bmatrix} \frac{3}{4} & 0 \\ \frac{1}{2} & \frac{1}{4} \end{bmatrix}.$

Thus $\qquad P_1^n \sim e\,V_1'$ where $V_1' = (V_1', 0')$, $= \left(\frac{4}{7}, 0\ \frac{3}{7}, 0\right)$.

In other words, for all I as n → ∞

$$p_{i1}^{(n)} \to \frac{4}{7}$$

$$p_{i3}^{(n)} \to \frac{3}{7}$$

$$p_{i2}^{(n)} \to 0$$

$$p_{i4}^{(n)} \to 0 .$$

Example 10: Stochastic Inventory Model(Seneta, 1981)

Consider that a store stocks a certain item, the demand for which is given by

p_k = Pr { k demand of the item in a week },

p_k> 0, k = 0, 1, 2 and p_k = 0, k ≥ 3.

Stocks are replenished at weekends according to the policy: not to replenish if there is any stock in store and to obtain 2 new items if there is no stock. Let X_n be the number of items at the end of n^{th} week, just before week's replenishment, if any, and Pr $\{X_0 = 3\}$ = 1.

Then $\{X_0\ n\ \geq 0\}$ is a Markov chain with state space S = { 0, 1, 2, 3 } and t. p. m.

$$P = \begin{bmatrix} p_2 p_1 p_0 & 0 \\ p_1 + p_2 p_0 & 0 & 0 \\ p_2 p_1 p_0 & 0 \\ 0 & p_2 p_1 p_0 \end{bmatrix} . \qquad p_0 + p_1 + p_2 =$$

1 .

The Markov chain is reducible, with a single closed class C with states 0, 1 and 2, the states being persistent non – null and aperiodic. The t. p. m (submatrix) is

$$P_1 = \begin{bmatrix} p_2 p_1 p_0 \\ p_1 + p_2 p_0 & 0 \\ p_2 p_1 p_0 \end{bmatrix}$$

The state 3 is transient.

Thus $$P = \begin{bmatrix} P_1 & 0 \\ R_1 & M \end{bmatrix}$$

Where $R_1 = (0,\ p_2,\ p_1)$, M = (p_0) and

$$0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Using Theorem 2.10 we get, $V_1' = \left(v_0,\ v_{1,},\ v_2\right)$

$$P_1^n \rightarrow e\, V_1'$$

$$V_1' P_1 = \ V_1' \text{ . } V_1' e = 1$$

$$p_{ij}^{(n)} \rightarrow v_j \text{ , j = 0, 1, 2.}$$

$$v_0 = (1 - p_0\,)^2 \,/\, \text{c}$$

$$v_1 = p_1 \,/\, \text{c}$$

$$v_2 = p_0\,(1 - p_0\,) \,/\, \text{c}$$

where $\quad C = \ (1 - p_0 + \ p_1)$.

Further $p^n \rightarrow e\, V_1'$ where $V' = (v_0,\ v_1,\ v_2, 0)$.

UNIT III

.MARKOV PROCESS WITH DISCRETE STATE SPACE

3.1 Introduction

Discrete state space Markov Processes has many applicants in day to day processes, such as inventory control in business, queuing systems and reliability theory. Poisson process is a versatile process which represents almost all random processes whose values move on a discrete space. The inter success time or inter-arrival time between two notified events are assumed to e exponential with parameter λ.

3.2. Poisson processes

Poisson is a special kind of Markov process with exponential inter arrival time. It is a stochastic process in continuous time with discrete state space which plays a vital role in modelling real life systems.

Description: Consider a random event ε such as incoming telephone calls, arrival of customer for services, occurrence of accidents etc.

Let us denote $N(t)$ the number of occurrence of the event ε in an interval of duration t. That is $N(t) denote$ the number of events ε occurred up to time epoch t. Then $\{N(t): t \geq 0\}$ is a counting process with time space R^+.

The path diagram of the process has step structure.

Fig 3.1

Let $p_{n(t)} = \Pr\{N(t) = n\}$. This probability is a function of time t and $\sum_{n=0}^{\infty} p_n(t) = 1$, where $\{p_n(t)\}$ represent probability distribution of the random variable $N(t)$ for every value of t.

The family of random variables, $\{N(t): t \geq 0\}$ is a stochastic process. Now we proceed to show that $N(t)$ follows a Poisson distribution with parameter λ, the mean is λt. Hence the stochastic process, $\{N(t)$: t≥ 0 } is a Poisson process.

3.2.1Poisson process and its Postulates:

1. **Independence:** The random variable, $(t + h) - N(t)$, the number of occurrences in the interval $(t, t + h)$ is independent of the number of occurrences prior to that interval.

2. **Homogeneity in time:** $p_n(t)$ depends only on the length t of the interval and is independent of the position of the interval. That is $p_n(t) =$ Pr {number of occurrence of event E in the interval $(t_{1,} t_1 + t)$ }

3. **Regularity:** In an interval of infinitesimal length h, the probability of exactly one occurrence is $\lambda\text{h} + \text{o(h)}$ and that of more than one occurrence is $o(h)$.

(Here o(h) is defined as $\lim_{h \to 0} \frac{o(h)}{h} = 0$.)

In other words, if the interval between t and t + h is of very short duration h, then

$$p_1(h) = \lambda\, h + o\,(h)$$

$$\sum_{k=2}^{\infty} p_k(h) = 0\,(h).$$

Since , $\sum_{n=0}^{\infty} p_n(h) = 1$,It follows that

$$p_0(h) = 1 - \lambda\, h + o\,(h)$$

(1.6)

Theorem 3.1. Under the postulates 1, 2 & 3, the random variable N (t) follows Position distribution with mean λ t. That is $p_n\,(t)$ is given by the Position law:

$$p_n\,(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n\,!}, \qquad n = 0, 1, 2, 3....$$

(1.7)

Proof:

Consider $p_n(t + h\ )$ for n ≥ 0:

The n events by epoch t + h can happen in the following mutually exclusive events

$$A_1,\ A_2,\ A_3, \ldots\ldots, A_{n+1}\ .$$

For n ≥ 1

A_1: n occurrences by epoch t and no occurrence event between t and t + h;

We have, $\Pr(A_1)$ = Pr $\{N\ (t) = n\}$Pr $\{N\ (h) = 0 \mid N(t) = n\}$ (1.8)

$= p_n(t)p_0(h)$

$= p_n(t)(1 - \lambda h)$ + o (h)

A_2: $(n-1)$ *occurrences* by t and 1 occurrences between t and t + h;

We have, $\Pr(A_2)$ = Pr $\{N\ (t) = n - 1\}$Pr $\{N\ (h) = 1 \mid N(t) = n - 1\}$

$= p_{n-1}(t)p_1(h)$

(1.9)

$= p_{n-1}(t)(\lambda h)$ + o (h)

For n ≥ 2

A_3: $(n-2)$ occurrences by epoch t and 2 occurrences between t and t + h;

We have, Pr $(A_3) = p_{n-2}(t)\{p_2(h)\} \leq\ p_2\,(h)$,

Same result holds for Pr (A_4), Pr (A_5),…..

Thus we have

$$\sum_{k=2}^{n} Pr\{A_{k+1}\} \leq \sum_{k=2}^{n} p_k(h) = \text{o}(h)$$

and so $p_n(t + h) = p_n(t)(1 - \lambda h) + p_{n-1}(t)(\lambda\ h)$+ o$(h)$, n ≥ 1

or, $\frac{p_n(t+h) = p_n(t)}{h}$= - λ$p_n(t)$ + λ$p_{n-1}(t)$ + $\frac{0(h)}{h}$

(1.10)

taking limit, as h → 0, we get

$$p'_n(t) = -\lambda[p_n(t) -\ p_{n-1}(t)],\ n \geq 1. \qquad (1.11)$$

For n = 0, we get

$$p_0(t+h) = p_0\,(t)\,p_0\,(h) = p_0(t)(1-\lambda h) + o(h)$$

or $$\frac{p_0(t+h) - p_0\,(t)}{h} = -\lambda\, p_0(t) + \frac{0(h)}{h}$$

whence, as h → 0, $p_0'(t) = -\lambda p_0(t)$. (1.12)

Initial condition:

Suppose that the process starts from scratch at time 0, so that N (0) = 0, i. e.

$p_0(0) = 1$; $p_n(0)$ for n ≠ 0. Z (1.13)

The differential – difference equations(1.11)and the differential equation(1.12) together with (1.13) completely specify the system. Their solutions give the probability distribution { $p_n(t)$ $of\ N\ (t)$. The solutions are given by

$$p_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \text{ n = 0, 1, 2,…..} \quad \square$$ (1.14

Proofs: (alternative)

We indicate here two other methods of solving these equations.

(1) The method of induction: The solution of (1.12) is given by $p_0(t) = Ce^{-\lambda t}$. Since $p_0(0) = 1$, we have C = 1 so that $p_0(t) = e^{-\lambda t}$. Consider (1.11) for n=1. Substituting the value of p_0 and solving the equation and using (1.13), we find $p_1(t) = \lambda$ t. $e^{-\lambda t}$. Thus (1.14) is seen to hold for n = 0 and 1. Assuming that it holds for (n-1) it can be shown likewise that it holds for any n. Hence, by induction, we get (1.14) for all n.

(2) The method generating function: Define the probability generating function

$$\mathrm{P}(s,t) = \sum_{n=0}^{\infty} p_n(t)\, s^n = \sum_{n=0}^{\infty} Pr\,\{N\,(t) = n\} s^n = \mathrm{E}\{s^{N(t)}\}\ .$$ (1.15)

Now $$\mathrm{P}(s,0) = \sum_{n=0}^{\infty} p_n(0)\, s^n = p_0(0) + p_1(0)\ \text{s} + \ldots = 1.$$ (1.16)

We have $\frac{\partial}{\partial t}\mathrm{P}(s,t) = \sum_{n=0}^{\infty}\frac{\partial}{\partial t}p_n(t)s^n = \sum_{n=1}^{\infty}p'(t)\,s^n + p_n'(t)s^n + p_0'(t);$

$\sum_{n=1}^{\infty}p_n(t)s^n = \mathrm{P}\,(s,t) - p_1(t);$

and $\sum_{n=1}^{\infty}p_{n-1}(t)s^n = \mathrm{s}\,\mathrm{P}(s,t).$ (1.17)

Multiplying (1.11) by s^n and adding over for n = 1, 2, 3,…. and using (1.17), we get

$$\frac{\partial}{\partial t}\mathrm{P}(s,t) - p_0'(t) = -\lambda\,[\{P\,(s,t) - p_0(t)\} - sP(s,t)]$$

or $\frac{\partial}{\partial t}\mathrm{P}(s,t) + \lambda p_0(t) = \mathrm{P}\,(s,t)\{\lambda\,(s-1)\} + \lambda p_0(t).$

Thus $\frac{\partial}{\partial t}\mathrm{P}(s,t) = \mathrm{P}(s,t)\{\lambda(\lambda\,st)\}.$ (1.18)

Solving (1.18), we get

$$\mathrm{P}\,(s,st) = \mathrm{A}\,e^{\lambda\,(s-1)t} \qquad (1.19)$$

Now P $(s,0)$ = 1 from (1.16), so that A = 1.

Hence the p. g. f of Position process is given by

$$\mathrm{P}\,(s,t) = e^{\lambda\,t\,(s-1)}$$

$$= e^{-\lambda\,t}\left\{\sum_{n=0}^{\infty}\frac{(\lambda\,s\,t)^n}{n\,!}\right\},$$

So that

$p_n(t) \equiv$ coefficient of s^n in P (s,t)

$$= \frac{e^{-\lambda\,t\,(\lambda\,t)^n}}{n\,!}, \qquad n \geq 0. \quad \square$$

Corollary 3.2.:

For a Poisson process, we have

(i) E $\{N\,(t)\} = \lambda$ t

and (ii) var $\{N\,(t)\} = \lambda$ t .

Proof: The proof is immediate consequence of previous theorem because $p_n(t)$ is well defined □

3.2.2Remark:

The mean number of occurrences in an interval of length t is λ t, so that the mean number of occurrences per unit time $(t = 1)$, i. e. in an interval of unit length is λ. The mean rate λ per unit time is known as the parameter of the Poisson process.

The mean and the variance of N (t) are function of t; in fact, its distribution is functionally dependent on t. As such the process $\{N\ (t), t\ \geq 0\}$ is not stationary – it is evolutionary

While $\{N\ (t), t\ \geq 0\}$ is a continuous parameter stochastic process with discrete state space, E$\{N\ (t)\}$ is a non – random continuous function of t.

Corollary 3.3:

If E occurred r times up to initial instant 0 from which t is measured, then the initial condition will be

$p_r(0)$ = 1, $p_n(0)$ = 0, n≠r.

Then $p_n(t)$ = Pr { Number N (t) of occurrences by epoch t is n-r, n ≥ r }

$$= \frac{e^{-\lambda t\,(\lambda t)^{n-r}}}{(n-r)\,!} \quad , \qquad n \geq r$$

$$= 0, \qquad n \geq r.$$

Example 1.

Suppose that customers arrive at a Bank according to a Poisson process with a mean rate of a minute. Then the number of customers N (t) arriving in an interval of duration t minutes follows Poisson distribution with mean at. If the rate of arrival is 3 per minute, then in an arrival of 2 minute, the probability that the number of customers arriving is:

(i) exactly 4 is

$$\frac{e^{-6}(6)^4}{4\,!} = 0.133,$$

(ii) greater than is

$$\sum_{k=5}^{\infty} \frac{e^{-6}(6)^k}{k\,!} = 0.714$$

(iii) less than 4 is

$$\sum_{k=0}^{3} \frac{e^{-6}(6)^k}{k\,!} = 0.152,$$

(using tables of Poisson distribution).

Example 2:

A machine goes out of order whenever a component part fails. The failure of this part is in accordance with a Poisson process with mean rate of 1 per week.

Then the probability that two weeks have elapsed since the last failure is e^{-2} = 0.135, being the probability that time t = 2 weeks, the number of occurrences is 0.

Suppose that there are 5 spare parts of the component in an inventory and that the next supply is not due in 10 weeks. The probability that the machine will not be out of order in the next 10 weeks is given by

$\sum_{k=0}^{5} \frac{e^{-10}(10)^k}{k\,!}$ = 0. 068,

Being the probability that the number of failures in t = 10 weeks will be than or equal to 5.

Example 3:

Estimation of the parameter of Poisson process. For a Poisson process $\{N\ (t)\}$, as t $\rightarrow \infty$

$$\Pr \rightarrow \left\{\left|\frac{N(t)}{t} - \lambda\right| \geq \varepsilon \right\}\ 0,$$

Where ε> 0 is a preassigned number.

This can be proved by applying Tshebyshev's lemma (for a r. v. X)

$$\Pr \{|X - E\ (X)| \geq a\} \leq \frac{var\ (X)}{a^2}, \text{ for } a > 0.$$

From the above, we have, for X = N (t),

$$\Pr\ \{|N(t) - \lambda\, t\ | \geq a\} \leq \frac{\lambda\, t}{a^2}$$

or $$\Pr \left\{\left|\frac{N(t)}{t} - \lambda\right| \geq \frac{a}{t}\right\} \leq \frac{\lambda\, t}{a^2}$$

or $$\Pr \left\{\left|\frac{N(t)}{t} - \lambda\right| \geq \varepsilon \right\} \leq \frac{\lambda}{t\varepsilon^2}\ .$$

Hence $$\Pr \left\{\left|\frac{N(t)}{t} - \lambda\right| \geq \varepsilon \right\} \rightarrow 0 \text{ as } t \rightarrow \infty$$

This implies that for large t, the observation N (t) / t may be used as a reasonable estimate of the mean rate λ of the process $\{N\ (t)\}$.

3.3POISSON PROCESS AND RELATED DISTRIBUTIONS

3. 3. 1Inter-arrival Time

With a Poisson process, $\{N(t), t \geq 0\}$, where $N(t)$ denotes the number of occurrences of an event E by epoch t, there is associated a random variable – the interval X between two successive occurrences of E. We proceed to show that X has a negative exponential distribution.

Theorem 3.3.2

The interval between two successive occurrences of a Poisson process $\{N(t), t \geq 0\}$ having parameter λ has a negative exponential distribution with mean $\frac{1}{\lambda}$.

Proof:

Let X be the random variable representing the interval between two successive occurrences of $\{N(t), t \geq 0\}$and let $Pr(X \leq x) = F(x)$ be its distribution function.

Let us denote two successive events by E_i and E_{i+1} and suppose that E_i occurred at the instant t_i. Then

$Pr(X > x) = Pr$ { E_{i+1} did not occur in $(t_i, t_i + x)$ given that E_i occurred at the instant t_i }

$= Pr$ { E_{i+1} did not occur in $(t_i, t_i + x) \mid N(t_i) = i$ }

(because of the postulate of independence)

= Pr { no occurrence takes place in an interval $(t_i, t_i + x)$ of length $x \mid N(t_i) = i$ }

= Pr { $N(x) = 0 \mid N(t_i) = i$ }$p_0(x) = e^{-\lambda x}, x > 0$.

Since i is arbitrary, we have for the interval X between any two successive occurrences,

$$F(x) = \Pr\{X \leq x\} = 1 - \Pr\{X > x\} = 1 - e^{-\lambda x}, x > 0.$$

The density function is

$$f(x) = F'(x) = \lambda e^{-\lambda x}, (x > 0).$$

It can be further proved that if X_i denote the interval between E_i and $E_{i+1}, i = 1, 2, \ldots.,$ then $X_1 X_2, \ldots$ also independent. We omit the proof which is outside the scope of this book. We state the result as follows:

Theorem 3.3.4

The intervals between successive occurrences of a Poisson process are identically and independently distributed random variable which follow the negative exponential law with mean $\frac{1}{\lambda}$. The convers also holds; this is given in Theorem 3.4 below. These two theorems give a characterization of the Poisson process.

Theorem 3.3.5

If the intervals between successive occurrences of an event E are independently distributed with a common exponential distribution with mean $\frac{1}{\lambda}$, then the events E form a Poisson process with mean λt.

Proof:

Let Z_n denote the interval between $(n-1)^{th}$ and n^{th} occurrence of a process $\{N(t), t \geq 0\}$ and let the sequence $Z_1, Z_2, \ldots$ be independently and identically distributed random variables having negative exponential distribution with mean $\frac{1}{\lambda}$. The sum $W_n = Z_1 + \ldots + Z_n$ is the waiting time up to the n^{th} occurrence, i. e. the time from the origin to the n^{th} subsequent occurrence. W_n has a gamma distribution with parameters λ, n. The p. d. f. $g(x)$ and distribution function F_{w_n} are given respectively by

$$g(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}, x > 0$$

And $$F_{w_n}(\mathrm{t}) = \Pr\{W_n \leq \mathrm{t}\} = \int_0^t g(x)dx.$$

The events $\{N(t) < n\}$ and $W_n = + \cdots + Z_n > t$ are equivalent. Hence the distribution functions $F_N(t)$ and F_{w_n} satisfy the relation

$$F_{w_n}(\mathrm{t}) = \Pr\{W_n \leq \mathrm{t}\} = 1 - \Pr\{W_n \leq \mathrm{t}\}$$

$$= 1 - Pr\{N(t) < n\} = 1 - Pr\{N(t) \leq (n-1)\}$$

$$= 1 - F_{N(t)}(n-1).$$

Hence the distribution function of $N(t)$ is given by

$$F_{N(t)}(n-1) = 1 - F_{w_n}(t)$$

$$= 1 - \int_0^t \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)} dx$$

$$= 1 - \frac{1}{\Gamma(n)} \int_0^{\lambda t} y^{n-1} e^{-y} \; dy$$

$$= \frac{1}{\Gamma(n)} \int_{\lambda t}^{\infty} y^{n-1} e^{-y} \; dy$$

$$= \sum_{j=0}^{n-1} \frac{e^{-\lambda t}(\lambda t)^j}{j\,!} \text{(integrating by parts).}$$

Thus the probability law of $N(t)$ is

$$p_n(t) = Pr\{N(t) = n\} = F_{N(t)}(n) - F_{N(t)}(n-1)$$

$$= \sum_{j=0}^{n} \frac{e^{-\lambda t}(\lambda t)^j}{j\,!} - \sum_{j=0}^{n-1} \frac{e^{-\lambda i}(\lambda t)^j}{j\,!}$$

$$= \frac{e^{-\lambda t}(\lambda t)^n}{n\,!} \;, n = 0, 1, 2, \ldots$$

Thus the process $\{N(t), t \geq 0\}$ is a Poisson process with mean λt. Note that Poisson process has independent exponentially distributed inter-arrival times and gamma distributed waiting times.

Example 7:

Suppose customers arrive at a service counter in accordance with a Poisson process with mean rate of 2 per minute (λ = 2 / minute). Then the interval between any two successive arrivals follows exponential distribution with mean $\frac{1}{\lambda} = \frac{1}{2}$ minute. The probability that the interval between two successive arrivals is

(i) more than 1 minute is

$$\Pr(X > 1) = e^{-2} = 0.135$$

(ii) 4 minutes or less

$$\Pr(X \leq 4) = 1 - e^{-4\times 2} = 1 - e^{-8} = 0.99967$$

(iii) between 1 and 2 minute is

$$\Pr(1 \leq X \leq 2) = \int_1^2 2\, e^{-2x} dx = e^{-2} e^{-4} =$$
0.0179.

Example 8:

Suppose that customers arrive at a service counter independently from two different sources. Arrivals occur in accordance with a Poisson process with mean rate of λ per hour from the first source and μ per hour from the second source. Since arrivals at the counter constitute a Poisson process with mean $(\lambda + \mu)$ per hour, the interval between any two successive arrivals has a negative exponential distribution with mean $\{1/(\lambda + \mu)\}$ hours.

For example, if taxis arrive at a spot from the north at the rate of 1 per minute and from the south at the rate 2 per minute in accordance with two independent Poisson process, the interval between arrival of two taxis has a (negative) exponential distribution with mean $\frac{1}{3}$ minute; the probability that a long person will have to wait more than a given time t can be found.

Poisson type of occurrences are also called purely random events and the Poisson process is called a purely random process. The reason for this is that the occurrence is equally likely to happen anywhere in $[0, T]$ given that only one occurrence has taken place in the interval. We state this by the following theorem.

Theorem 3.3.6:

Given that only one occurrence of a Poisson process $N(t)$ has occurred by epoch T, then the distribution of the time interval γ in $[0, T]$ in which it occurred is uniform in $[0, T]$, i. e.

$$\Pr\{\, t < \gamma \le t + dt \mid N(T) = 1 \,\} = \frac{dt}{T}, \qquad 0 < t < T$$

Proof: We have $\Pr\{\, t < \gamma \le t + dt \,\} = \lambda e^{-\lambda t}\, dt,$

$$\Pr\{\, N(T) = 1\} = e^{-\lambda T}(\lambda T),$$

and $\Pr\{\, N(T) = 1 \mid \gamma = t\} = e^{-\lambda(T-t)}$

The last one being the probability that there was no occurrence in the interval of length (T-t). Hence

$$\Pr\{\, t < \gamma < t + dt \mid N(T) = 1\}$$

$$= \frac{\Pr\{\, t < \gamma < t + dt \text{ and } N(T) = 1\}}{Pr\{N(T) = 1\}}$$

$$= \frac{\Pr\{\, t < \gamma < t + dt \,\}\Pr\{\, N(T) = 1 \mid \gamma = t\}}{Pr\{N(T) = 1\}}$$

$$\lambda e^{-\lambda t}\, dt e^{-\lambda T}(\lambda T) = dt/T.$$

It follows that $\Pr\{\, \gamma \le s \mid \; N(T) = 1\} = \frac{s}{T}$

It may be said that a Poisson process distributes points at random over the infinite interval $[0, \infty]$ in the same way as the uniform distributes points at random over a finite interval $[a, b]$. □

3.4Properties of Poisson Process

3.4.1. Additive property:

Sum of two independent Poisson processes is a Poisson process. Let $N_1(t)$and $N_2(t)$ be two Poisson processes with parameters λ_1 ,λ_2 respectively and let

$$N(t) = N_1(t) + N_2(t).$$

The p. g. f. of $N_i(t)$ (I = 1, 2) is

$$E\left\{s^{N_i(t)}\right\} = e^{\lambda_i(s-1)t}$$

The p. g. f. of $N(t)$ is

$$E\left\{s^{N(t)}\right\} = E\left\{s^{N_1(t)+N_2(t)}\right\}$$

And because of independent of $N_1(t)$ and $N_2(t)$, we have

$$E\left\{s^{N(t)}\right\} = E\left\{s^{N_1(t)}\right\} E\left\{s^{N_2(t)}\right\}$$

$$= \left\{e^{\lambda_1(s-1)t}\right\}\left\{e^{\lambda_2(s-1)t}\right\}$$

$$= e^{(\lambda_1+\lambda_2)(s-1)t}$$

Thus $N(t)$ is a Poisson process with parameter $\lambda_1 + \lambda_2$.

The result can also be proved as follows:

$$\Pr\{N(t) = n\} = \sum_{r=0}^{n} \Pr\{N_1(t) = r\}. \Pr\{N_2(t) = n - r\}$$

$$= \sum_{r=0}^{n} \frac{e^{-\lambda_1 t}(\lambda_1 t)^r}{r\,!} \cdot \frac{e^{-\lambda_2 t}(\lambda_2 t)^{n-r}}{(n-r)\,!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)^t}\{(\lambda_1+\lambda_2)t\}^n}{n\,!}, \qquad n \geq 0. \ \square$$

Hence N (t) is a Poisson process with parameter$(\lambda_1 + \lambda_2)$.

3.3.2. **Difference of two independent Poisson processes**: The probability distribution of $N(t) = N_1(t) - N_2(t)$ is given by,

$$\Pr\{N(t) = n\} = e^{-(\lambda_1+\lambda_2)t}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{n}{2}} I_{|n|}\left(2t\sqrt{\lambda_1\lambda_2}\right), \quad n = 0, \pm 1, \pm 2, \ldots,$$

Where $I_n(x) = \sum_{r=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2r+n}}{r!\,\Gamma(r+n+1)}$

Is the modified Bessel function of order n (≥ -1) .

Proof: (i) The p. g. f. of N (t) is

$$\text{E}\left\{s^{N(t)}\right\} = \text{E}\left\{s^{N_1(t) - N_2(t)}\right\}$$

$$= \text{E}\left\{s^{N_1(t)}\right\} \text{E}\left\{s^{-N_2(t)}\right\},$$

Because of the independence of $N_1(t)$ and $-N_2(t)$. Thus

$$\text{E}\left\{s^{N(t)}\right\} = \text{E}\left\{s^{N_1(t)}\right\} \text{E}\left\{\frac{1}{s}^{N_2(t)}\right\}$$

$$= \exp\{\lambda_1 t (s-1)\} \exp\{\lambda_1 t (s^{-1}-1)\}$$

$$= \exp\left\{-(\lambda_1+\lambda_2)t\right\} \exp\left\{\lambda_1 ts + \lambda_2 \frac{t}{s}\right\}.$$

Pr $\{N(t) = n\}$ is given by the coefficient of s^n in the expansion of the right hand side of as a series in positive and negative powers of s.

(ii) Pr$\{N(t) = n\}$can also be obtained directly as follows:

$$\Pr\{N(t) = n\} = \sum_{r=0}^{\infty} Pr\{N_1(t) = n+r\}\, Pr\{N_2(t) = r\}$$

$$= \sum_{r=0}^{\infty} \frac{e^{-\lambda_1 t}(\lambda_1 t)^{n+r}}{(n+r)!} \frac{e^{-\lambda_2 t}(\lambda_2 t)^r}{r!}$$

$$= e^{-(\lambda_1+\lambda_2)t}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{n}{2}} \sum_{r=0}^{\infty} \frac{\left(t\sqrt{\lambda_2\lambda_2}\right)^{2r+n}}{r!\,(r+n)!}$$

$$= e^{-(\lambda_1+\lambda_2)t}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{n}{2}} I_{|n|}\left(2\,t\sqrt{\lambda_2\lambda_2}\right).$$

It may be noted that

(1) the difference of two independent Poisson processes is not a Poisson process:

(2) $I_{-n}(t) = I_{|n|}(t)$, = 1, 2, 3,…..:

(3) the first two moments of $N(t)$ are given by

E $\{N(t)\} = (\lambda_1 + \lambda_2)\, t$ and E $\{N^2(t)\} = (\lambda_1 + \lambda_2)$ t + $(\lambda_1 + \lambda_2)^2 t^2$.

Example 4.

If passengers arrive at a taxi stand in accordance with a Poisson process with parameter λ_1 and taxis arrive in accordance with a Poisson process with parameter λ_2 then $N(t)$ = N_1 (t) $-N_2$ (t) gives the excess of passengers over taxis in an interval t. The distribution of (t) , i. e., Pr $\{N(t) = n\}$, n = 0, $\pm$ 1, $\pm$ 2…. Is given by *. The mean of $N(t)$ is $(\lambda_1 + \lambda_2)$ t, which is > = 0 < 0 according as

$$\lambda_1 >= \lambda_2 \text{ ; and var } \{N(t)\} = (\lambda_1 + \lambda_2)\, t$$

3.3.3. Decomposition of a Poisson process:

A random selection from a Poisson process yields a Poisson process. Suppose that $N(t)$, the number of occurrences of an event E in an interval of length t is a Poisson process with parameter λ. Suppose also that each occurrence of E has a constant probability p of being recorded, and that the recording of an occurrence is independent of that of other occurrences and also of $N(t)$.

Theorem: 3.3.4

If $M(t)$ is the number of occurrences recorded in an interval of length t, then $M(t)$ is also a Poisson process with parameterλp.

Proof:

The event $\{M(t)\}$ can happen in the following mutually exclusive ways:

A_r : E occurs $(n + r)$ times by epoch t and exactly n out of $(n + r)$ occurrences are recorded, probability of each occurrence recorded being p, (r = 0, 1, 2,…).

We have

Pr (A_r) = Pr $\{E\ occurs\ (n + r)\ times\ by\ epoch\ t\}$. Pr $\{n\ occurrences\ are\ recorded\ given\ that\ the\ number\ of\ occurrences\ is\ n = r$

$$\frac{e^{-\lambda t}(\lambda t)^{n+r}}{(n+r)!}\binom{n+r}{n} p^n q^r \ .$$

Hence Pr $\{M(t) = n\}$ = $\sum_{r=0}^{\infty} Pr(A_r)$

$$= \sum_{r=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{n+r}}{(n+r)!} \binom{n+r}{n} p^n q^r$$

$$= e^{-\lambda t} \sum_{r=0}^{\infty} \frac{(\lambda pt)^n (\lambda qt)^r}{n!\, r!}$$

$$= e^{-\lambda t} \frac{(\lambda pt)^n}{n!} \sum_{r=0}^{\infty} \frac{(\lambda qt)^r}{r!}$$

$$= e^{-\lambda t} \frac{(\lambda pt)^n}{n!} e^{\lambda q t} = \frac{e^{-\lambda pt}(\lambda pt)^n}{n!} \quad \square$$

3.3.5 Remark: We can interpret the above as follows

For a Poisson process $\{N(t)\}$, the probability of an occurrence in an infinitesimal interval h is proportional to the length h, the proportionality being λ. Now for $\{M(t)\}$, the probability of a recording in the interval h is proportional to the length h, the constant of proportionality being λp. Thus $\{M(t), t \geq 0\}$ is a Poisson process with parameter λp.

3.3.6. Continuation of property 3

The number $M_1(t)$ of occurrences not recorded is also a Poisson process with parameter λq = $\lambda(1-p)$ and $M(t)$ and $M_1(t)$ are independent.

Thus by random selection a Poisson process $\{N(t), t \geq 0\}$ of parameter λ is decomposed into two independent Poisson process $\{M(t), t \geq 0\}$ and $\{M_1(t), t \geq 0\}$ with parameters λp and $\lambda\ (1-p)$ Respectively.

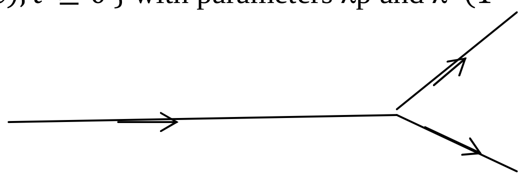


Fig. 3.2 Decomposition rates

As an example, suppose that the births occur in accordance with a Poisson process with parameter λ. If the probability that an individual born is male is p, then the male births form a Poisson process with parameter λp and the female births form an independent Poisson process with parameter $\lambda(1-p)$.

More generally, a Poisson process $\{N(t)\}$ with parameter λ may be decomposed into r stream of Poisson processes. If $p_1 + \ldots + p_r = 1$ then the Poisson process is decomposed into r independent Poisson process with parameters $\lambda p_1, \lambda p_2, \ldots, \lambda p_r$.

Theorem 3.3.7.(Poisson process and binomial distribution)

If $\{N(t)\}$ is a Poisson process and s < 1, then

$$\Pr\{N(s) = \mathrm{k} \mid N(t) = \mathrm{n}\} = \binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\left(\frac{s}{t}\right)\right)^{n-k}.$$

Proof:

$$\Pr\{N(s) = \mathrm{k} \mid N(t) = \mathrm{n}\} = \frac{\Pr\{N(s)=\mathrm{k} \text{ and } N(t)=\mathrm{n}\}}{\Pr\{N(t)=\mathrm{n}\}}$$

$$= \frac{\Pr\{N(s)=\mathrm{k} \text{ and } N(t-s)=\mathrm{n-k}\}}{\Pr\{N(t)=\mathrm{n}\}}$$

$$= \frac{\Pr\{N(s)=\mathrm{k}\}\Pr\{N(t-s)=\mathrm{n-k}\}}{\Pr\{N(t)=\mathrm{n}\}}$$

$$= \frac{e^{-\lambda s}(\lambda s)^k}{k!} \cdot \frac{e^{-\lambda(t-s)}[\lambda(t-s)]^{n-k}}{(n-k)!} \Big/ \frac{e^{-\lambda s}(\lambda t)^n}{n!}$$

$$= \frac{n!}{k!(n-k)!}\frac{s^k(t-s)^{n-k}}{t^n}$$

$$= \binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\left(\frac{s}{t}\right)\right)^{n-k} \square$$

Theorem 3.3.8:

If $\{N(t)\}$ is a Poisson process then the auto – correlation coefficient between $N(t)$ and $N(t+s)$ is $\{t/(t+s)\}^{\frac{1}{2}}$.

Proof:

Let λ be the parameter of the process; then

E $\{N(T)\}$ = λT, var $\{N(T)\}$ = λT,

and E $\{N^2(T)\}$ = λT + $(\lambda\mathrm{T})^2$ for T = t and t+s.

Since $N(t)$ and $\{N(t+s) - N(t)\}$ and independent, $\{N(t), t \geq 0\}$ being a Poisson process,

E$\{N(t)N(t+s)\}$ = E$[N(t)\{N(t+s) - N(t) +$ $\mathcal{N}t$

= E $[N(t)N(t)]$+ $[N(t)\{N(t+$ s−$\mathcal{N}t$

$$= \text{E}\ \{N^2(t)\} + \text{E}\ \{N\ (t)\}\ \text{E}\{N(t + s-N t.$$

Hence $\qquad \text{E}\ \{N(t)N\ (t+s)\} = (\lambda t + \lambda^2 t^2) + \lambda t.\ \lambda s.$

Thus the auto-covariance between $N(t)\ and N\ (t+s)$ is given by

$$\text{C}(t, t+s) = \text{E}\ \{N(t)N\ (t+s)\} - \text{E}\ \{N\ (t)\}\ \text{E}\{N(t + s$$

$$= (\lambda t + \lambda^2 t^2 + \lambda^2 ts) - \lambda t(\lambda t + \lambda s) = \lambda t.$$

Hence the autocorrelation function

$$\rho(t, t+s) = \frac{C(t,t+s)}{\{var\ N(t)\ var\ N(t+s)\}^{\frac{1}{2}}}$$

$$= \left\{\left(t/(t+s)\right)\right\}^{\frac{1}{2}}.$$

It can be shown that

$$\rho(t, t') = \frac{\min(t,t')}{(tt')^{\frac{1}{2}}}.$$

This is the autocorrelation function of the process.

Theorem 3.3.9

If $\{N(t)\}$ is a Poisson process with parameter λ, then $\{N(t) - \lambda t, t\ \geq 0\}$ is a continuous parameter martingale.

Example 5:

A radioactive source emits particles at a rate of 5 per minute in accordance with a Poisson process. Each particle emitted has a probability 0. 6 of being recorded. The number $N(t)$ of particles recorded in an interval of length t is a Poisson process with rate 5 × 0. 6 = 3 per minute, i. e. with mean 3t and variance 3t. In a 4 – minute interval the probability that the number of particles recorded is 10 is equal to $e^{-12}(12)^{10}/10\ !$ = 0.104.

Example 6:

A person enlists subscriptions to a magazine, the number enlisted being given by a Poisson process with mean rate 6 per day. Subscribers may subscribe for 1 or 2 years independently of one another with respective probabilities $\frac{2}{3}$ and $\frac{1}{3}$. The number of subscribers $N(t)$ enrolled by the person in time t days is a Poisson process with mean rate 6t: the number $N_1(t)$

enrolled for 1 – year period is a Poisson process with mean $\left(6 \times \frac{2}{3}\right) t$ = 4t and the number $N_2(t)$ enrolled for 2 – year period is a Poisson process with mean $\left(6 \times \frac{1}{3}\right) t$ = 2t .

If the commission received is a for 1 – year subscription and b for a 2 – year subscribtion, then the total commission earned in period t is given by

$$X(t) = aN_1(t) + b\, N_2(t)$$

We have $$E\,\{X(t)\} = aE\{N_1(t)\} + bE\{N_2(t)\}$$

$= 4at + 2bt$

And $$\text{var}\, X(t) = 4a^2t + 2b^2t.$$

3.3.10 Note: The process $\{X(t), t > 0\}$ is a compound Poisson process

3.5 GENERALISATIONS OF POISSONPROCESS

There are several directions in which the Poisson process discussed in the previous section can be generalized. We consider some of them here.

3.5.1 Poisson Process in Higher Dimensions

We have considered so far the one – dimensional case: the occurrences take place at random instants of time t (say, $t_1, \ t_2, ...$) and thus we were concerned with distribution of points on a line. Instead, we may have the two – dimensional case.

Consider the two – dimensional case, such that for the number $N(\Delta a)$ of occurrences in an element of area Δa, we have, for infinitesimal Δa,

$$\Pr\{N(\Delta a) = 1\} = \lambda\Delta a + o(\Delta a),$$

$$\Pr\{N(\Delta a) = k\} = o(\Delta a), k \geq 2$$

and $$\Pr\{N(\Delta a) = 0\} = 1 - \lambda\Delta a + o(\Delta a),$$

Thus, if the number of occurrences in non – overlapping areas are mutually independent, the number $N(a)$ of occurrences in an area a will be a Poisson process with mean λa. Here in place of one – dimensional t, we consider two – dimensional a. Similarly, we can describe Poisson process in higher dimensions.

3.5. 2 Poisson Cluster Process (Compound or Cumulative Poisson Process)

Discrete Case

We considered in that only one event can occur at an instant of occurrence. Now let us suppose that several events can happen simultaneously at such an instant, i. e. we have a cluster(of occurrences) at a point. We assume that:

(i) The number $N(t)$ of clusters in time t, i. e. the points at which clusters occur constitute a Poisson process with mean rate λ.

(ii) Each cluster has a random number of occurrences, i. e. the number X_i of occurrences in i^{th} cluster is a. r. v. The various numbers of occurrences in the different clusters are mutually independent and follow the same probability distribution:

$$\Pr\{X_i = k\} = P_k,\ k = 1, 2, 3, ...$$

$$i = 1, 2, 3, \ldots ..$$

Having p. g. f. $P(s) = \sum_{k=1}^{\infty} p_k s^k$.

Theorem 3.5.3

If $M(t)$ denotes the total number of occurrences in an interval of length t under the conditions (i) and (ii) stated above, then the generating function of $M(t)$ is given by

$$G\big(P(s)\big) = \exp\,[\lambda t\,\{P(s) - 1\}].$$

Proof:$M(t)$ is the sum of a random number of terms, i. e.

$$M(t) = \sum_{i=1}^{N(t)} X_i\ ,$$

Where If $N(t)$ is a Poisson process with meanλt.

Now $P(s)$ is the p. g. f. of X_i and $G(s)$ is the p. g. f. of (t) . Thus,

$$G(s) = exp\{\lambda t(s - 1)\}.$$

Hence by the p. g. f. of $M(t)$ is given by

$$G(Ps) = exp\{\lambda t P(s) - 1\}.$$

3.5.4 Note:

(I) $M(t)$ is called a compound Poisson process It is to be noted that $M(t)$ is not necessarily Poisson. Poisson cluster process arise in bulk queues, where customers arrive or are served in groups.

(ii) A compound Poisson process is not a point process; it is what is called a jump process.

(iii) Suppose that claims against a company occur in accordance with aPoisson process with mean λt, and that individual claims X_i are i. i. d. with distribution $\{p_k\}$, then $M(t)$ represents the total at epoch t. If a represents initial reserve and c the rate of increase of the reserves in the absence of claims, then the total reserve at epoch t is $A + ct - M(t)$, and negative reserve implies 'ruin'.

3.5.5Continuous Case

Now suppose that non – negative variables X_i are continuous having d. f. $F(x) = \Pr\{X_i \leq x\}$,p. d. f. f(x), and L. T.

$$f^*(s) = \int_0^\infty e^{-sx} f(x) dx$$

Then, it can be shown (as before) that the L. T. of $X(t)$ is given by

$$E\{\exp(-sX(t))\} = \exp\,[\lambda t\{f^*(s) - 1\}]$$

$X(t)$ is known as a continuous compound Poisson process. By compound Poisson process, we shall generally mean (discrete) compound Poisson process.

The L. T. of $\{X(t) > x\}$ is given by

$$\int_0^\infty e^{-sx} \Pr\{X(t) > x\}\, dx = \frac{1 - \exp\{\lambda t(f^*(s) - 1)\}}{s}$$

3.5.6. Compound Poisson process and linear combination of independent Poisson processes:

Consider Example 6.

The process $X(t) = aN_1(t) + bN_2(t)$ is a linear combination of two independent Poisson processes. The process can also be expressed as the compound Poisson process

$$X(t) = \sum_{i=1}^{N(t)} X_i$$

where X_i, the amount of commission received form a subscription, is a random variable such that

$$\Pr(X_i = a) = \frac{2}{3}\ and\ \Pr\{X_i = b\} = \frac{1}{3}\ .$$

If follows that

And $var\{X(t)\} = 6t\ E\{X_i^2\} = 6t.\frac{1}{3}.(2a^2 + b^2) = 4a^2t + 2b^2t.$

The two approaches are equivalent. The result may be stated in a more general form as follows:

Let $a_k > 0, k = 1, 2, \dots, r(\geq 2)\ and\ \Pr(X_i = a_k) = p_k\ for\ each\ i, \sum pk = 1.$

Then $X(t) = \sum_{i=1}^{N(t)} X_i$, where $\{N(t)\}$ is a Poisson process with parameter λ, is a compound Poisson process. For $t > 0, let\ N_i\ (t)$ be the number of jumps of value a_i for the process $\{X(t)\}$ which occur prior to t, Then we have

$$X(t) = a_1N_1(t) + \cdots + a_rN_r\ (t)$$

,where $\{N_k(t), t \geq 0\}$ is a Poisson process with parameter $\lambda p_k\ and\ N_1(t), \dots, N_r\ (t)$ are mutually independent. $N(t)$ is decomposed into r independent Poisson processes $N_k(t), k = 1, 2, \dots., r.$

Example 9.

Customers arrive at a store in groups consisting of 1 or 2 individuals with equal probability and the arrival of groups is in accordance with a Poisson process with mean rate λ.

Here $p_k = \Pr\{X_i = k\} = \frac{1}{2}(for\ k = 1, 2)$

$$= 0(otherwise);$$

Hence $P(s) = \sum_k p_k s^k = \frac{1}{2}s + \frac{1}{2}s^2$

And so the generating function of $M(t)$, the total number of customers arriving in time t is

$$G(s) = \exp\left\{\lambda t\left[\frac{1}{2}(s + s^2) - 1\right]\right\}.$$

The mean number of customers arriving in time t equals $E\{X_i\}\{\lambda t\} = \frac{3}{2}\lambda t.\ For \lambda = \frac{1}{2}$per minute and t=4 minutes, the generating function will be

$$\exp\left[2\left\{\frac{1}{2}(s + s^2) - 1\right\}\right] = \{\exp(-2)\}\{\exp\ (s + s^2)\},$$

And the probability that the total number of arrivals is exactly 4 is

$$e^{-2}\left(\frac{1}{4\,!}+\frac{1}{2\,!}+\frac{1}{2\,!}\right)=0.141.$$

Example 10.

Suppose that the number of arrival epochs in an interval of length t is given by a Poisson process $\{N\,(t), t \geq 0\}$ with mean a and that the number of units arriving at an arrival epoch is given by a zero – truncated Poisson variable$X_i, i = 1, 2, \ldots$ with parameterλ. Then the total number $M(t)$ of units which arrive in an interval of length t is a Poisson cluster process with p. g. f.

$$G\big(P(s)\big) = \exp\,[at\{P(s) - 1\}]$$

Where P(s) is the p. g. f. of zero – truncated Poisson process namely

$$P(s) = (exp\lambda - 1)^{-1}(\exp(\lambda s) - 1).$$

Hence $$G\big(P(s)\big) = \exp\left[at\left\{\frac{(\exp(\lambda s)-1)}{(exp\lambda-1)} - 1\right\}\right].$$

Example 11. (An application in inventory theory)

Suppose that X_i are i. i. d. decapitated geometric r. $v.'\,s$ such that

$$Pr\{X_i = k\} = q^{k-1}p, k = 1, 2, 3, \ldots\,,\ where\ \ p + q = 1$$

Then the generating function of X_i is , $G(s) = ps/(1 - qs)$,

and $\big(P(s)\big) = \exp[\lambda t\{ps/(1 - qs) - 1\}]$.

From the function G(s), we get$p_0 \equiv Pr\{M(t) = 0\} = e^{-\lambda t}$

$$p_k \equiv Pr\{M(t) = k\} is\ given\ by$$

$$p_k = \frac{p\lambda t}{k}\sum_{j=1}^{k} q^{j-1} j\, p_{k-j}, \qquad k \geq 1.$$

3.5.7. Remark: Several authors have used this as a model for *lead time* demand of a commodity. It has been shown that this distribution fits actual data for demand of units of an EOQ or consumable type inventory item during stock *replenishment* or *lead time.*

UNIT IV

BIRTH AND DEATH PROCESS AND CONINUOUS TIME MARKOV CHAIN

4.1 Introduction

A stochastic process whose state space moves back and forth by unit measure in state space is called Birth-Death process. A simple example for birth –death process is the queuing system in which arrival customer to the counter is a *birth* and the service completion in a server is equivalent to *death* event. Inventory control system with one for –one ordering policy is also an example for Birth- Death process. In this unit we study the pure birth and pure death process together with Birth-Death process.

4.2 Birth –Death Process

4.2.1 Pure Birth- Death Process:

First we considera pure birth process, where Pr { Number of births between t and t+h is k, given the number of individuals at epoch t is n}

Is given by
$$p(k,h\,|n,t) = \begin{cases} \lambda_n h + o(h), & k = 1 \\ o(h), & k \geq 2 \\ 1 - \lambda_n h + o(h), & k = 0. \end{cases} \qquad (4.1)$$

The above holds for all $n \geq 0; \lambda_0$ may or not be equal to zero. Here k is a non – negative integer which implies that there can only be an increase by k, i. e. only births are considered possible. Now we suppose that there could also be a decrease by k, i. e. death(s) is also considered possible. In this case we shall further assume that

Pr { Number of births between t and t+h is k, given the number of individuals at epoch t is n}

Is given by $$q(k, h \mid n, t) = \begin{cases} \mu_n h + o(h), & k = 1 \\ o(h), & k \geq 2 \\ 1 - \mu_n h + o(h), & k = 0. \end{cases}$$ (4. 2)

The above holds for $n \geq k$; further$\mu_0 = 0$,Which is known as a birth and death process. Through a birth there is an increase by one and through a death, there is a decrease by one in the number of "individuals". The probability of more than one birth or more than one death in an interval of length h is $o(h)$. Let $N(t)$ denotes the total number of individuals at epoch t starting from$t = 0$ and let$p_n(t) = Pr\{N(t) = n\}$.Consider the interval between 0 and $t + h$; suppose that it is split into two periods $(0, t) and\ [\,t, t + h]$. The event $\{N(t + h) = n, n \geq 1\}$, (having probability $p_n(t + h)$ can occur in a number of mutually exclusive ways.

These would include events involving more than one birth and / or more than one death between t and $+h$. By our assumption, the probability of such an event is $o(h)$. There will remain four other events to be considered:

A_{ij} : $(n - i + j)$ individuals by epoch t, i birth and j death between t and $t + h, i, j = 0, 1$.

We have $$Pr(A_{00}) = p_n(t)\{1 - \lambda_n h + o(h)\}\{1 - \mu_n h + o(h)$$

$$= p_n(t)\{1 - (\lambda_n + \mu_n)h + o(h)\};$$

$$Pr(A_{10}) = p_{n-1}(t)\{\lambda_{n-1} h + o(h)\}\{1 - \mu_{n-1} h + o(h)$$

$$= p_{n-1}(t)\lambda_{n-1} h + o(h);$$

$$Pr(A_{01}) = p_{n-1}(t)\{1 - \lambda_{n+1} h + o(h)\}\{\mu_{n+1} h + o(h)$$

$$= p_{n+1}(t)\mu_{n+1} h + o(h);$$

and $$Pr(A_{11}) = p_n(t)\{\lambda_n h + o(h)\}\{\mu_n h + o(h)\}$$

$$= o(h).$$

Hence we have, for n ≥ 1

$$p_n(t + h) = p_n(t)\{1 - (\lambda_n + \mu_n) + h\} + p_{n-1}(t)\lambda_{n-1} h + p_{n+1}(t)\mu_{n+1} h + o(h) \quad (4.3\ )$$

Or $\frac{p_n(t+h)-p_n(t)}{h} = -(\lambda_n+\mu_n)p_n(t) + \lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1}(t) + \frac{o(h)}{h}$.

And taking limits, as $h \to 0$, we have

$$p_n'(t) = -(\lambda_n+\mu_n)p_n(t) + \lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1}(t),\ n \geq 1. \qquad (4.4)$$

For $n = 0$, we have

$$p_0(t+h) = p_0(t)\{1-(\lambda_0 h + o(h)\} + p_1(t)\{1-(\lambda_0 h + o(h)\}\{\mu_n h + o(h) \qquad (4.5)$$

$$= p_0(t) - \lambda_0 h p_0(t) + \mu_n h p_1(t)$$

or $\frac{p_0\,(t+h)-\,p_0(t)}{h} = -\lambda_0 p_0(t) + \mu_1 p_1(t) + \frac{o(h)}{h}$

Taken limit as h→ 0, we get

$$p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t) \qquad (4.6)$$

If at epoch$t = 0$, there were $i(\geq 0)$ individuals, then the initial condition is

$$p_n(0) = 0,\ n \neq i, p_i(0) = 1. \qquad (4.7)$$

The above equations (4. 4) and (4. 6) are the *equations of the birth and death process*. The birth and death processes play an important role in queuing theory. They also have interesting applications in diverse other fields such as economics, biology, reliability theory etc.

4.2.2 REMARKS:

The result about existence of solutions of (4. 4) and (4. 6) is stated below without proof. For srbitrary

$\lambda_n \geq 0, \mu_n \geq 0$, there always exists a solution $p_n(t)(\geq 0)$ such that $\sum p_n(t) \leq 1$. If λ_n, μ_n are bounded, the solution is unique and satisfies $\sum p_n(t) = 1$.

4. 2. 3 BIRTH AND DEATH RATES

Some particular values of $\lambda_n\ and\ \mu_n$ are of special interest. When $\lambda_n = \lambda, i.e.\ \lambda_n$ is independent of the population size n, then the increase may be thought of as due to an external source such as immigration. When $\lambda_n = n\,\lambda$, we have case of (linear) birth; $\lambda_n h = n\lambda h$ may by considered as the probability of one birth in an interval of length h given that n individuals are present (at the instant from which the interval commences)the probability of one individual giving a birth being λh, (i. e. rate of birth in unit interval is λ per individual). Here $\lambda_0 = 0$.

When$\mu_n = \mu$, the decrease may effected due to the emigration factor.When,$\mu_n = n\mu$, we have the case of death, the rate of death in unit interval being μ per individual**.**

Particular Cases

I. Immigration – Emigration Process

For $\lambda_n = \lambda \; and \; \mu_n = \mu$ we have what is known as immigration – emigration process. The process associated with the simple queuing model $M/M/1$ in such a process.

II. Linear Growth Process

(a)Generating Function:

IN the Yule –Furry process one is concerned with a population whose members can give birth only but cannot die. Let us consider the case where both births and deaths can occur. Suppose that the probability that a member gives birth to a new member in a small interval of length h is $\lambda h + o(h)$ and the probability that a member dies is $\mu h + o(h)$. Then, if n members are present at the instant t, the probability of one birth between t and $t + h$ is $n\lambda h + o(h)$ and that of one death is $\mu h + o(h), n \geq 1$.

We have thus a birth and death process with

$$\lambda_n = n\lambda, \mu_n = \mu (n \geq 1), \lambda_0 = \mu_0 = 0.$$

If $X(t)$ denotes the total number of members at time t, then from (4. 4) and (4. 6) we have the following differential – difference equations for $p_0(t) = \Pr\{X(t) = n\}$:

$$p_n'(t) = -n(\lambda + \mu)p_n(t) + \lambda(n-1)p_{n-1}(t) + \mu(n+1)p_{n+1}(t), \; n \geq 1$$
(4. 8)

$$p_n'(t) = \mu p_1(t).$$

If the initial population size is i, i. e. $X(0) = i$, then we have the initial condition $p_i(0) = 1$ and $p_n(0) = 0, n \neq i$.

Let $P(s,t) = \sum_{n=1}^{\infty} p_n(t)s^n$ be the p. g. f. of $\{p_n(t)\}$.

Then $\frac{\partial P}{\partial s} = \sum_{n=1}^{\infty} \mathrm{n} p_n(t) s^{n-1}$ and $\frac{\partial P}{\partial t} = \sum_{n=0}^{\infty} p_n'(t) s^n$.

Multiplying (4. 8) by s^n and adding over $n = 1, 2, 3, ...$ and adding (4. 9) thereto, we get

$$\frac{\partial P}{\partial t} = -(\lambda + \mu) \sum_{n=1}^{\infty} n p_n(t) s^n + \lambda \sum_{n=1}^{\infty} (\mathrm{n} -$$

1) $pn{-}1tsn$

$$= +\mu\{\sum_{n=1}^{\infty} (\mathrm{n} + 1) p_{n+1}(t) s^n + p_n(t)\}$$

$$= -(\lambda + \mu) s \frac{\partial P}{\partial s} + \lambda s^2 \frac{\partial P}{\partial s} + \mu \frac{\partial P}{\partial s}$$

$$= \{\boldsymbol{\mu} - (\lambda + \mu)\boldsymbol{s} + \lambda s^2\} \frac{\partial P}{\partial s}.$$

$P(s, t)$ thus satisfies a partial differential equation of Lagrangian type. We shall not discuss here the method of solution; the solution with the initial condition $X(0)i$. Is given by

$$P(s,t) = \left[\frac{\mu(1-s) - (\mu - \lambda s)\, e^{-(\lambda-\mu)t}}{\lambda(1-s) - (\mu - \lambda s) e^{-(\lambda-\mu)t}}\right]^i$$

$$= \left[\frac{\mu\{1 - e^{-(\lambda-\mu)t}\} - \{\mu - \lambda e^{-(\lambda-\mu)t}\}s}{\{\lambda - \mu\, e^{-(\lambda-\mu)t}\} - \lambda\,\{1 - e^{-(\lambda-\mu)t}\}s}\right]^i$$

(4. 11)

Explicit expression for $p_n(t)$ can be obtained from the above by expanding $P(s, t)$ as a power series in s.

(b) Mean Population Size:

We can obtain the mean population size by differentiating$P(s, t)$ partially with respect to s and putting s=1. It can however be obtained directly form (4. 8) and (4. 9) without obtaining $P(s, t)$ as follows:

Let $E\{X(t)\} = M(t) = \sum_{n=1}^{\infty} \mathrm{n} p_n(t)$

and $E\{X^2(t)\} = M_2(t) = \sum_{n=1}^{\infty} \mathrm{n}^2 p_n(t)$.

Multiplying both sides of (4. 8) by n adding over for $n = 1, 2, 3, ...$, we have

$$\sum_{n=1}^{\infty} \mathrm{n} p_n'(t) = -(\lambda + \mu) \sum_{n=1}^{\infty} \mathrm{n}^2 p_n(t) + \lambda \sum_{n=1}^{\infty} \mathrm{n}(\mathrm{n} - 1)\, p_{n-1}(t) + \mu \sum_{n=1}^{\infty} n(\mathrm{n} + 1) p_{n+1}(t) . \quad (4.\ 12)$$

Now $\sum_{n=1}^{\infty} \mathrm{n}(\mathrm{n} - 1)\, p_{n-1}(t) = \sum_{n=1}^{\infty} (\mathrm{n} - 1)^2\, p_{n-1}(t) + \sum_{n=1}^{\infty} (\mathrm{n} - 1)\, p_{n-1}(t)$

$$= M_2(t) + M(t);$$

$$\sum_{n=1}^{\infty} \text{n}(\text{n}+1)\, p_{n+1}(t) = \sum_{n=1}^{\infty} (\text{n}+1)^2\, p_{n+1}(t) - \sum_{n=1}^{\infty} (\text{n}+1)\, p_{n+1}(t)$$

$$= \{M_2(t) - p_1(t)\} - \{M(t) - p_1(t)\}$$

$$= M_2(t) - M\,(t);$$

And $\sum_{n=1}^{\infty} \text{n}p'(t) = M'(t)$.

Hence from (4. 12) we get

$$M'(t) = -(\lambda + \mu\,)\, M_2(t) + \lambda\{M_2(t) + M(t)\} + \mu\{M_2(t) - M(t)\}$$

$$= (\lambda + \mu\,) M(t).$$

The solution of the above differential equation (that M(t) satisfies) is easily found to be

$$M\,(t) = \text{C}e^{(\lambda-\mu)t}$$

The initial condition gives$M(0) = \sum_{n=1}^{\infty} np_n(0) = i, whence$ C$= M(o) = i.$

We have therefore,

$$M(t) = ie^{(\lambda-\mu)t}\ .$$

The second moment $M_2(t)\ of\ X(t)$ can also be calculated in the same way.

Limiting case:

As $t \to \infty$, the mean population size $M(t)$ tends to 0 for $\lambda < \mu$ (birth rate smaller than death rate) or to ∞ for $\lambda < \mu$ (birth rate greater than death rate) and to the constant value I when $\lambda = \mu$.

(c) Extinction Probability:

Since $\lambda_0 = 0$, 0 is an absorbing state, i. e. once the population size reaches 0, it remains at 0 thereafter. This is the interesting case of extinction of the population. We can determine the probability of extinction as follows:

Suppose, for simplicity, that $X(0) = 1, i.e.$ the process starts with only one member at time 0. Then from (4.11) we can write $P(s,t)$ as

$$P(s,t) = \frac{a-bs}{c-ds} = \frac{a}{c}.\frac{1-bs/a}{1-ds/c}$$

Where $a = \mu\left\{1 - e^{-(\lambda-\mu)t}\right\}$

And $c = \lambda - \mu e^{-(\lambda-\mu)t}$

$Pr\{X(t) = 0\} = p_0(t)$, the constant term in the expression of $P(s,t)$ as a power series in s, is given by

$$\frac{a}{c} = \frac{\mu(1-e^{-(\lambda-\mu t)})}{\lambda-\mu e^{(\lambda-\mu)t}}.$$

The probability that the population will eventually die out is given by lim $p_0(t)$ as $t \to \infty$ and can be obtained from the above by letting $t \to \infty$.

If $\lambda > \mu$, then $$lim_{t\to\infty}\, p_0(t) = lim_{t\to\infty} \frac{\mu(1-e^{-(\lambda-\mu)t})}{\lambda-\mu e^{-(\lambda-\mu)t}}$$

$$= \frac{\mu}{\lambda} < 1.$$

If $\lambda < \mu$, then $$lim_{t\to\infty}\, p_0(t) = lim_{t\to\infty} \frac{\mu\{1-e^{-(\lambda-\mu)t}\}}{\lambda-\mu e^{-(\mu-\lambda)t}} = 1.$$

And $$lim_{t\to\infty}\, p_n(t) = 0 \text{ for } n \neq 0.$$

In other words, the probability of ultimate extinction is 1 when $\mu < \lambda$ and is $\frac{\mu}{\lambda} < 1$ when $\mu < \lambda$.

III**. Linear Growth Immigration**

In II, we have $\lambda_0 = 0$ and, as a result, if the population size reaches zero at any time, it remains at zero thereafter. Here 0 is an absorbing state. If we consider $\lambda_n = n\lambda + \alpha(\alpha > 0), \mu_n = n\mu(n \geq 0)$ we get what is known as a linear growth process with immigration, where 0 is not an sbsorting state.

IV**. Immigration – Death Process**

If $\lambda_n = \lambda$ and $\mu_n = n\mu$, we get what is known as an immigration – death process. This corresponds to the Markovian queue with infinite number of channels, i. e. the queue $M/M/\infty$.

V. **Pure Death Process**

Here $\lambda_n = 0$ for all n, i. e. an individual cannot give birth to a new individual and the probability of death of an individual in $(t, t+h)$is $\mu h + o(h)$. Then , if n individuals are present at time t , the probability of one death in $(t, t+h)$ is $n\mu h + O(h)$.

The birth and death process is a special case of continuous time Markov process with discrete state space $S = \{0,1,2, \dots \dots\}$ such that the probability of transtition from i to j in Δt time is $(\Delta t\ )$ whenever

$|i-j| \geq 2$. In other words changes takes place through transitions only from a state to its immediate neighboring state.

4.3Continuous Time Markov Chains:

4.3.1 Definition

A continuous time parameter MARKOV process $\{X(t): t \geq 0\}$ with discrete state space $N = \{0,1,2 \ldots\ldots\}$ is considered for this section. Assume that $\{X(t): t \geq o\}$ *is a time* homogeneous Markov chain.

So the probability of a transition from state I to state j during the time interval $(T, T+t)$ does not depend on the initial time T, but depends only on the elapsed time t and on the initial and terminal states I and j. We can thus write

$$Pr\{X(T+t) = j \mid X(T) = i\} = p_{ij}(t), i,\ j = 0, 1, 2, \ldots, \quad t \geq 0.$$

In particular, we write $\quad Pr\{X(t) = j \mid X(0) = i\} = p_{ij}(t),$

From the definition of transition probability distribution, we have$0 \leq p_{ij}(t) \leq 1$ for each $i, j, t,$

and$\sum_j p_{ij}(t) = 1.$

Let$p_j(t) = Pr\{X(t) = j\}$ be the state probability at epoch j, then

$$p_j(t) = Pr\{X(t) = j\}$$

$$= \sum_i Pr\{X(t) = j \text{ } and \text{ } X(0) = i\}$$

$$= \sum_i Pr\{X(0) = i\} Pr\{X(t) = j \mid X(0) = i\}$$

$$= \sum_i Pr\{X(0) = i\} p_{ij}(t).$$

Now we have $\sum_j p_j(t) = 1$ for $t \geq 0$.

Let us denote the transition probability matrix of the Markov Chain by

$$P(t) = \left(p_{ij}(t)\right).$$

Setting$p_{ij}(0) = \delta_{ij}$, we get, $P(0) = I$. Also assume here that the functions $p_{ij}(t)$ are continuous and differentiable for $t \geq 0$.

4.3.2 The waiting time for a change of state:

Suppose that $\{X(t): t \geq o$ is a homogeneous Markov process and that at time$t_0 = 0$,the state of the process $X(t_0) = X(0) = i$ is known. The time taken for a change of state I is a random variable, say τ. This random time period τ is called the waiting time to reach a different state from state i.

We have $Pr\{\tau > s + t \mid X(0) = i\}$

$= Pr\{\tau > s + t \mid X(0) = i, \tau > s\} Pr\{\tau > s \mid X(0) = i\}$

$= Pr\{\tau > s + t \mid X(s) = i,\} Pr\{\tau > s \mid X(0) = i\}.$

If we denote $\overline{F}(t) = Pr\{\tau > t \mid X(0) = i\}, t > 0$ then the above can be written as

$$\overline{F}(s + t) = \overline{F}(s)\overline{F}(t), \ for\ s, t > 0.$$

The above relation is satisfied if $\overline{F}(t)$ is the form$e^{-\lambda t}$, $t > 0$, $\lambda > 0$.

Thus the waiting time τ has an exponential distribution with parameter λ, which is called the transition density from state i. The distribution function is the same for all i.

4.3.3 Chapman – Kolmogorov Equations:

The transition probability $p_{ij}(t + T)$ is the probability that the given state was I at epoch 0, it is in state j at epoch $t + T$; but in passing from state i to state j in time $(t + T)$ the process moves through some state k in time t, Thus

$$p_{ij}(t + T) = \sum_i Pr\{X(t + T) = j, \quad X(t) = k \mid X(0) = i\}$$

$$= \sum_i Pr\{X(0) = i,\ X(t) = k,\ X(t + T) = j\} / \Pr\{X(0) = i\}$$

$$= \sum_k \frac{\{Pr\{X(0)=i,\ X(t)=k\}\}}{\Pr\{X(0)=i\}} \times \frac{Pr\{X(0)=i,\ X(t)=k,\ X(t+T)=j\}}{\Pr\{X(0)=i,\ X(t)=k\}}$$

$$= \sum_i Pr\{X(t) = k \mid X(0) = i\} \Pr\{X(t + T) = j \mid X(0)=i,\ X(t)=k\}.$$

Since $\{X(t): t \geq 0\}$ is a Markov process,

$$\Pr\{X(t + T) = j \mid X(0) = i,\ X(t) = k\}$$

$= \Pr\{ X(t+T) = j \mid X(t) = k\}$

$= p_{ij}(T).$

Hence we get the probabilistic relation

$p_{ij}(t+T) = \sum_i p_{ik}(t)\, p_{kj}(T),$ for all states i, j and $t \geq o, T \geq 0.$ (4.5)

Above equation (4.20) is called Chapman- Kolmogorov equation.

4.3.4 Remark:

We can also write Chapman –Kolmogorov equation in matrix form:

$P(t+T) = P(t).P(T).$ (4.6)

The equations 4.5 and 4.6 are also equivalent to the relation we already proved an the case of Discrete Markov Chains.

Denote the right – hand derivative at zero by

$$a_{ij} = \frac{d}{dt} p_{ij}(t)|_{t=0} \;;\; i \neq j$$
(4. 7)

$$a_{ij} = \frac{d}{dt} p_{ii}(t)|_{t=0} \,.$$

Then $a_{ij} = lim_{\Delta t \to 0} \frac{p_{ij}(\Delta t) - p_{ij}(0)}{\Delta t} = lim_{\Delta t \to 0} \frac{p_{ij}(\Delta t)}{\Delta t}$

or $p_{ij}(\Delta t) = a_{ij}(\Delta t) + o(\Delta t),\; i \neq j$
(4.8)

and $a_{ii} = lim_{\Delta t \to 0} \frac{p_{ii}(\Delta t) - p_{ii}(0)}{\Delta t} = lim_{\Delta t \to 0} \frac{p_{ii}(\Delta t) - 1}{\Delta t}$

or $p_{ij}(\Delta t) = 1 + a_{ij}(\Delta t) + o(\Delta t).$
(4.9)

It can be seen from the above relations that $a_{ij} \geq 0, i \neq j$ and $a_{ii} < 0.$

From $\sum_j p_{ij}(t) = 1,$ using (4.7), we get

$$\sum_j a_{ij} = 0$$

$$\sum_{j \neq i} a_{ij} = -a_{ii}.$$
(4.10)

The quantities a_{ij} are called transition densities and the matrix

$$A = (a_{ij})$$

Is called the transition density matrix or rate matrix of the process. The matrix is such that

(1) its off – diagonal elements are non- negative and the diagonal elements are negative;

(2) the sum of the elements of each row is zero, the sum of the off – diagonal elements being equal in magnitude but opposite in sign to the diagonal elements being elements.

Differentiating (4.5) with respects to T, we get

$$p'_{ij}(t+T) = \frac{\partial}{\partial T} p_{ij}(t+T) = \sum_k p_{ik}(t) \frac{d}{dT} p_{kj}(T).$$

Putting T=0, we get $\quad p'_{ij}(t) = \sum_k p_{ik}(t) a_{kj}.$ (4.11)

Or, in matrix notation $P'(t) = P(t)A.$ (4.11a)

Similarly we can get $\frac{d}{dT} p_{ij}(T) = \sum_k p_{ik} p_{kj}(T).$

Replacing T by t, we can write this as

$$p'_{ij}(t) = \sum_k p_{ik} p_{kj}(t) \qquad (4.12)$$

or

$$P'(t) = AP(t). \qquad (4.12a)$$

Equations (4.11) and (4.12) which give Chapman – Kolmogorov equations as differential equations are called respectively Forward and Backward Kolmogorov equations.

Solution of the Equations for a Finite State Process

When the rate matrix is given, the equations (4.11) or (4.11a) together with the initial conditions $p_{ij} = \delta_{ij} (or\ P(0) = I)$ yield as solution the unknown probabilities $p_{ij}(t)$. We consider below a method of solution for a process with finite number of states. From (4.11a) we see at once that the solutions can be written in the form

$$P(t) = P(0)e^{At} = e^{At}$$

(4.13)

Where the matrix $$e^{AT} = I + \sum_{n=1}^{\infty} \frac{A^n t^n}{n!}$$

(4.13a)

Assume that eigenvalues of A are all distinct. Then from the spectral theorem of matrices, we have

$$A = HDH^{-1}$$

Where H is a non – singular matrix and D is the diagonal matrix having for its diagonal elements the eigenvalues of A. Now, 0 is an eigenvalue of A and if $d_i \neq 0, i = 1, \dots, m$ are the other distinct eigenvalues, then

$$D = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & d_i & \dots & 0 \\ \dots & & \dots & \dots \\ \dots & & \dots & \dots \\ 0 & 0 & \dots & d_m \end{bmatrix}.$$

We then have

$$D^n = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & d_1^n & \dots & 0 \\ \dots & & \dots & \dots \\ \dots & & \dots & \dots \\ 0 & 0 & \dots & d_m^n \end{bmatrix}$$

And $$A^n = HD^nH^{-1}$$

Substituting in (4.13), we get

$$P(T) = I + \sum_{n=1}^{\infty} \frac{(HD^nH^{-1})t^n}{n!}$$

$$= H\left\{I + \sum_{n=1}^{\infty} \frac{D^n t^n}{n!}\right\} H^{-1}$$

$$= He^{Dt}H^{-1}$$

Where $$e^{Dt} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & e^{d_1 t} & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & e^{d_m t} \end{bmatrix}$$

The right – hand side of (4.14) gives explicit solution of the matrix $P(t)$. Note that even in the general case when the eigenvalues of A are not necessarily distinct, a canonical representation of $A = SZS^{-1}$ exists. Using this, $P(t)$ can be obtained in a modified form.

Example 5(a): Poisson process:

If events occur in accordance with a Poisson process $N(t)$ with mean λt, then

$$p_{i,i+1}(\Delta t) = Pr\{\text{the process goes to state i+1 from state I in time } \Delta t\}$$

$$= \Pr\{\text{one event occurs in time } \Delta t\}$$

$$= \Pr\{N(\Delta t) = 1\}$$

$$= \lambda\Delta t + o(\Delta t),$$

$$p_{i,i}(\Delta t) = 1 - \lambda\Delta t + o(\Delta t)$$

and $\quad p_{i,j}(\Delta t) = o(\Delta t),\ j \neq i,\ i+1.$

By comparing with (4.8) and (4.9), we have

$$a_{i,i+1} = \lambda, a_{i,i} = -\lambda, a_{i,j} = 0 \; for \; j \neq i, i+1.$$

The rate matrix is $A = (a_{ij}) = \begin{bmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -\lambda & \lambda & \dots & 0 \\ \dots & & & \dots & \end{bmatrix}$

The Kolmogorov forward equations are

$$p'_{i,i}(t) = -\lambda p_{i,i}(t)$$

$$p'_{i,i}(t) = -\lambda p_{i,i}(t) + \lambda p_{i,i-1}(t), j = i+1, i+2, \dots \qquad (4.15)$$

Let $p_j(t) = \Pr\{N(t) = j\} \; and \; p_0(0) = 1, p_n(0) = 0, n \neq 0$. Using (4.2) we get $p_j(t) \equiv p_{0j}(t), j = 0,1,2 \dots$ Thus (4.15) become identical with (1.11) and (1.12) so that $p_j(t) = e^{-\lambda t}(\lambda t)^i / j\,!$ Similarly. With $p_{ij}(0) = 1, j = i, p_{ij}(0) = 0, i \neq j,$ we get

$$p_{ij}(t) = \frac{e^{-\lambda t}(\lambda t)^{j-i}}{(j-i)!}.$$

Example 5(b) Two – state process:

Suppose that a certain system can considered to be in two states: “Operating” and “Under repair”. Suppose that the lengths of operating period and period under repair are independent random variables having negative exponential distribution with means $\frac{1}{b}$ and $\frac{1}{a}$ respectively (a, b>0). The evolution of the system can be described by a Markov process with two states 0 and 1.

Now

$$p_{01}(\Delta t) = \Pr\{change\ of\ state\ from\ 0\ to\ 1\ in\ time \Delta t\}$$

$$= \Pr\{repair\ being\ completed\ in\ time \Delta t\}$$

$$= a\Delta t + o(\Delta t)$$

And $$p_{01}(\Delta t) = \Pr\{change\ of\ state\ from\ 1\ to\ 0\ in\ \Delta t\}$$

$$= b\Delta t + o(\Delta t).$$

Thus the transition densities are

$$a_{01} = a, \qquad a_{01} = b$$

And $$a_{00} = -a, \qquad a_{11} = -b$$

So that $$A = \begin{bmatrix} -a & a \\ b & -b \end{bmatrix}.$$

The Kolmogorov forward equations, for $i = 0, 1$, are

$$p'_{i0}(t) = -a\, p_{i0}(t) + b\, p_{i1}(t)$$

$$p'_{i1}(t) = a\, p_{t0}(t) - b\, p_{i1}(t).$$

Now we proceed to find the transition probabilities $p_{ij}(t)$.

Using $$p_{00}(t) + p_{01}(t) = 1, p_{10}(t) + p_{11}(t) = 1,$$

We get $$p'_{00}(t) = -(a+b)p_{00}(t) = b$$

And $$p'_{11}(t) = +(a+b)p_{11}(t) = a.$$

The solution of the first of these differential equation is

$$p_{00}(t) = \frac{b}{a+b} + Ce^{-(a+b)t}.$$

With $p_{00}(0) = 1, we\ get\ C = \frac{a}{a+b}$, so that

$$p_{00}(t) = \frac{b}{a+b} + \frac{a}{a+b}e^{-(a+b)t}.$$

Hence $$p_{01}(t) = 1 - p_{00}(t) = \frac{a}{a+b} - \frac{a}{a+b}e^{-(a+b)t}.$$

Similarly, the solution of the second differential equation with initial condition $p_{11}(0) = 1$, gives

$$p_{11}(t) = \frac{a}{a+b} - \frac{a}{a+b} e^{-(a+b)t} .$$

and hence $$p_{10}(t) = 1 - p_{11}(t)$$

$$= \frac{b}{a+b} - \frac{b}{a+b} e^{-(a+b)t}$$

Let $p_j(t)$ be the probability that the system is in state j at time $t, j = 0, 1, and\ let\ p_0(0) = 1,\ p_n(0) = 0, n \neq 0.$ Then $p_j(t) \equiv p_{0j}(t), j = 0, 1.$

4.3.5 Alternative Method:

We consider the above example to show how to proceed with matrix method of solution: this method is useful when Kolomogorov differential equations are easily solvable.

Here $A = \begin{bmatrix} -a & a \\ b & -b \end{bmatrix}$ has eigenvalues 0 and $-(a+b)$, corresponding right eigenvectors being $(1,1)'$ and

$(a-b)'$ respectively. The Kolmogorov forward equation

$$P'(t) = (p'_{ij}(t)) = P(t)A$$

Has as solution (as given in (4.14))

$$P(t) = He^{D(t)}H^{-1} \qquad (4.16)$$

Where $$H = \begin{bmatrix} 1 & a \\ 1 & -b \end{bmatrix}$$

And $$H^{-1} = \frac{1}{a+b} \begin{bmatrix} b & a \\ 1 & -1 \end{bmatrix}$$

The diagonal matrix D whose elements are the eigenvalues of A is

$$D = \begin{bmatrix} 0 & 0 \\ 0 & -(a+b) \end{bmatrix}$$

So that $$e^{D(t)} = \begin{bmatrix} 1 & 0 \\ 0 & e^{-(a+b)t} \end{bmatrix}.$$

Thus from (4.16)

$$P(t) = \begin{bmatrix} 1 & a \\ 1 & -b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{-(a+b)t} \end{bmatrix} \begin{bmatrix} b & a \\ 1 & -1 \end{bmatrix} / [(a+b)]$$

$$= \frac{1}{a+b}\begin{bmatrix} b + ae^{-(a+b)t} & a - a\,e^{-(a+b)t} \\ b - b\,e^{-(a+b)t} & a + be^{-(a+b)t} \end{bmatrix}$$

We have $p_{00}(t) = \frac{b}{a+b} + \frac{a}{a+b}e^{-(a+b)t}$, $p_{01}(t) = 1 - p_{00}(t)$

And $p_{11}(t) = \frac{a}{a+b} + \frac{b}{a+b}e^{-(a+b)t}$, $p_{10}(t) = 1 - p_{11}(t)$

4.3.6 Limiting Distribution (Ergodicity of Homogeneous Markov Process)

We recall the result on limiting distribution of certain types of Markov chains as given in Previous Theorems. we recall:

$$V(P - I) = 0, V = (v_0, v_1, \ldots.), Ve = 1 \quad (4.17)$$

A similar elegant result holds for continuous parameter Markov processes as well. We shall state the result without proof. Here similar definitions for the classification of the states will be used.

Theorem 4.3.7

Suppose that the time – homogeneous Markov process $\{X(t)\}$ is irreducible having aperiodic non – null persistent states; also that its t. p. m. is $P(t) = \left(p_{ij}(t)\right), i, j = 0, 1, 2, \ldots$ and the matrix of transition densities (or rate matrix) is

$$A = (a_{ij}),$$

Where $a_{ij} = p'_{ij}(t)|_{\,t=0}$.

Then given any state j, $lim_{t\to\infty}\, p_{ij}(t) = v_j$ (4.18)

Exists and is the same for all initial states $i = 0, 1, 2, \ldots$ The asymptotic values v_j represent a probability distribution, i. e. $0 \le v_j \le 1, \sum_j v_j = 1.$

The values v_j can then be determined as solutions of the system of linear equations

$$\sum_j v_j a_{ij} = 0.\ j = 0, 1, 2, \ldots \quad (4.19)$$

or in matrix notation, $VA = 0, V = (v_0, v_1, \ldots)$ (4.19a) by using the normalizing condition $ve = 1, that\ is, \sum_j v_j = 1$.

Note1: The eq. (4.19) can be obtained from the forward Kolmogorov eq. (5.11) by putting

$$lim_{t\to\infty} p_{ij}(t) = v_j\ , and\ lim_{t\to\infty} p'_{ij}(t) = 0.$$

Note2: The eqs. (4.17) and (4.19a) for discrete and continuous parameter processes respectively are similar in structure. The matrices $(P - I)$ and A both have non – negative off – diagonal elements, strictly negative diagonal elements and zero row sums. If the number of states are finite, say, m, then both $(P - I)$ and A are of rank $(m - 1)$. Then V can be easily determined from any of the $(m - 1)$ equations (out of m equations) contained in the relations (4.17) and (4.19a) and the normalizing condition $\sum_j v_j = 1$.

Example 5(c):

Consider the two – state process given Example 5(b). Here

$$v_0 = lim_{t\to\infty} p_{00}(t) = lim_{t\to\infty} p_{10}(t) = \frac{b}{(a+b)}$$

and $$v_1 = lim_{t\to\infty} p_{10}(t) = lim_{t\to\infty} p_{11}(t) = \frac{a}{(a+b)}\ .$$

These limiting probabilities can also be obtained from the equations (4.19) which become

$$-av_0 + bv_1 = 0$$

$$av_0 - bv_1 = 0.$$

Here there are two states and so the matrix A is of rank 1 and the two equations are equivalent. Each of them yields $v_1 = (a/b)v_0$. Using the normalizing condition $v_0 + v_1 = 1$, we get

$$v_0 = \frac{b}{(a+b)}, v_1 = \frac{a}{(a+b)}\ .$$

When the number of states is finite and when only the limiting probabilities v_j are needed it is easier and more convenient to determine them from (4.19) or (4.19a).

Example5(d):

Consider a M/M/m queuing system which has m service channels, the demand for service arises in accordance with a Poisson process with parameter

a. Suppose that the service time in each channel is exponentially distributed with parameter b. Further assume that there is no waiting space facility, in the sense that a demand which is received when all the m channels are busy is rejected and leaves the system. This system is called a Erlang ***loss system.***

. Describe a Markov process$\{X(t), t \geq 0\}$, where $X(t)$ denotes the number of busy channels at time t; it has (m+1) states 0, 1,…,m . Suppose that the system is in state $k, then$ it implies that k channels are busy. The transition probabilities of the Markov process are given by

$p_{i,i+1}(\Delta t) = Pr\{$one demand is received for processing in time $\Delta t\}$

$$= a\Delta t + o(\Delta t), \quad 0 \leq i \leq m$$

$$p_{i,j}(\Delta t) = o(\Delta t),\ j > i + 1.$$

$Pr\{$one service completion occurred in time Δt $\}$

$$= b\Delta t + o(\Delta t).$$

Suppose I channels are working$Pr\{$one service demand is met in time Δt $\}$ $= ib\Delta t + o(\Delta t)$,

i. e. $\quad p_{i,i-1}(\Delta t) = ib\Delta t + o(\Delta t)$

and $\quad p_{ij}(\Delta t) = o(\Delta t), j < i - 1.$

Thus $\quad a_{01} = a\ a_{00} = -a$

$$a_{ij} \begin{array}{ll} = ib, & j = i - 1 \\ = a\,, & j = i + 1 \\ = -\ (a + ib), & j = 1 \end{array} \Bigg\} 1 \leq i < m$$

$$a_{m,m-1} = mb, a_{m,m} = -mb$$

$$A = \begin{bmatrix} -a & a & 0 & 0 & \ldots & 0 & 0 \\ b & -(a+b) & a & 0 & \ldots & 0 & 0 \\ 0 & 2b & -(a+2b) & a & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots. & \ldots \\ 0 & 0 & 0 & 0 & \ldots & mb & -mb \end{bmatrix}$$

The equations (4.19) becomes

$$-av_0 + bv_1 = 0$$

$$av_{j-1} - (a + jb)v_j + (j + 1)bv_{j+1} = 0, \qquad j = 1, 2, \ldots, m - 1$$

$$av_{m-1} - bmv_m = 0.$$

The solution of these equations can be obtained recursively, from the equation

$$v_1 = \left(\frac{a}{b}\right) v_0.$$

Writing the second equation with $j = 1$ and putting there this value of v_1, we get$v_2 = \frac{1}{2}\left(\frac{a}{b}\right)^2 v_0$.

Proceeding in this way, we get

$$v_j = \left(\frac{1}{j\,!}\right)\left(\frac{a}{b}\right)^j v_0, \quad j = 0, 1, 2, \ldots, m.$$

From the total probability condition $\sum_{j=0}^{m} v_j = 1$, one gets $v_0 = \frac{1}{\left(\sum_{j=0}^{m}\left(\frac{1}{i\,!}\right)\left(\frac{a}{b}\right)^i\right)}$.

Formulas giving v_j are called ***Erlang's formulas***.

Note : The probability that a demand is rejected (lost) is given by

$$v_m = \frac{\frac{\left(\frac{a}{b}\right)^m}{m\,!}}{\frac{\sum_{j=0}^{m}\left(\frac{a}{b}\right)^i}{i\,!}}$$

This is known as ***Erlang's loss formula (Erlang's Blocking formula)*** and is denoted by $B(m, a/b)$. The formula is still being usedin telecommunication systems. Properties of the above Erlang Loss system has been studied by many researchers in the last five decades. The relation $v_m \equiv B(m, \frac{a}{b})$ can also be expressed as

$$[B(m, \tfrac{a}{b})]^{-1} = [\text{m/(a/b)}]\ [B(m-1, a/b]^{-1} + 1.$$

This recursive relation with the initial value $\left(0, \frac{a}{b}\right) = 1$ is used for computation of values of $v_m, m = 1, 2, \ldots$

Example 5(e): Machine Interference Problem

Consider that there are m identical machines. Each of the machines operates independently and is serviced a single servicing unit in case of break down. The operating time and serving time of each machine are independently distributed as exponential distribution with parameters $b\ and\ a$ respectively. Then the number of machines in operating condition at time t constitutes a Markov process$\{X(t), t \geq 0\}$ with state space $(0, 1, .., m)$.

Example5(f):

Two – channel service system. Consider Example 5(d) with $m = 2$. Suppose that the service channels are numbered I and II and suppose that we are interested in whether particular channels are busy (B) or free (F). Let the ordered pair(I, j) denote the state of the system, where i refers to that of the firstchannel and j to that of the second. The four states of the system$(F, F), (F, B), (B, F)$ and (B, B) may be denoted by 0, 1, 2, 3 respectively. The process $\{X(t)\}$ denoting that the states of the system in terms of the two channels may be described by a Markov process with state space 0, 1, 2, 3.

Assume that when both the channels are free a demand may join either of the channels for service with equal probability $\left(\frac{1}{2}\right)$. Thus, when both the channels are free, demands to each of the channels flow in accordance with a Poisson process with parameter $a/2$. We have

$$p_{0,j}(\Delta t) = \left(\frac{a}{2}\right)\Delta t + o(\Delta t), \qquad j = 1, 2$$

$$p_{1,j}(\Delta t) = b\Delta t + o(\Delta t), \qquad j = 0$$

$$= a\Delta t + o(\Delta t), \qquad j = 3$$

$$p_{2,j}(\Delta t) = b\Delta t + o(\Delta t), \qquad j = 0$$

$$= a\Delta t + o(\Delta t), \qquad j = 3$$

$$p_{3,j}(\Delta t) = b\Delta t + o(\Delta t), \qquad j = 1, 2$$

And for all other combinations of $i \neq j, p_{i,j}(\Delta t) = o(\Delta t)$.

Thus the matrix A will be

$$\begin{bmatrix} -a & \frac{a}{2}\,\frac{a}{2} & 0 & \\ b & -(a+b) & 0 & a \\ b & 0 & -(a+b) & a \\ 0 & b & b & -2b \end{bmatrix}$$

The normal equation (4.19) become

$$-av_0 \quad + bv_1 \quad + bv_2 \quad = 0$$

$$\left(\frac{a}{2}\right) v_0 \quad - (a+b)v_1 \quad + bv_3 = 0$$

$$\left(\frac{a}{2}\right) v_0 \quad - (a+b)v_1 \quad + bv_3 = 0$$

$$av_1 \qquad + av_2 \quad - 2bv_3 = 0.$$

From the second and third equations we get $v_1 = v_2$ and then from the first we get $v_0 = (2b/a)v_1$ and from the last $v_3 = (a/b)v_{1.}$ Utilising $\sum_{i=0}^{3} v_1 = 1,$ we at once get

$$v_1 = v_2 = \frac{ab}{a^2+2ab+2b^2}$$

$$v_0 = \frac{2b^2}{a^2+2ab+2b^2}$$

$$v_3 = \frac{a^2}{a^2+2ab+2b^2} \ .$$

BLOCK II

UNIT : V WEINER PROCESSES AND BRANCHING PROCESSES

5.1 Markov Process with Continuous State Space

Poisson processes is a real life process with continuous time with discrete counting state space. But in most of the real life problems Markov Problems have continuous state space. For example , level of water in a dam over a continuous time space, Life time of a electronic device over a continuous time are Stochastic processes with continuous state space. In mathematical term $\{X(t): t \epsilon T\}$where $T = (-\infty, \infty)$ and $X(t)\ \epsilon\ (-\infty, \infty)$ is a stochastic process with continuous time space and continuous state space.

5.2 Introduction: BROWNIAN MOTION PROBLEMS

Poisson process is a process in continuous time with a discrete state space. Here in a small interval of time Δt, there is either no change of state or there is only one change, the probability of more than one change being of the order of Δt. In this unit we shall consider Markov processes such that in an infinitesimal interval, there is a small change of state or displacement. In such a process, changes of state occur continually all the time and the state space is continuous. Because of the connection with the theory of diffusion, Markov processes with continuous state space are also known as diffusion processes. A particle under diffusion or undergoing Brownian motion is also known as a Brownian a fixed axis.

At epoch t, let $X(t)$ be the displacement along a fixed axis of a particle undergoing Brownian motion and let $X(0) = x_0$. Consider an interval (s, t) of time; let us regard this interval as the sum of a large number of small intervals. The total displacement$\{X(t) - X(s)\}$ in this interval can be regarded as the limit of the sum of random displacement over the small intervals of time. Suppose that the random displacements are independently distributed. Then it can be seen that the central – limit theorem applies, whence it follows that the total displacement $\{X(t) - X(s)\}$ is normally distributed. Further, suppose that the displacement $\{X(t) - X(s)\}$ depends on the length of the interval (s, t)

and not on the time – point s that $\{X(t) - X(s)\}$ has the same distribution as $\{X(t+h) - X(s+h)\}$ for all $h > 0$.

It is to be noted that here both time and space variables are continuous. The equations of the process obtained by taking limits of both time and space variables will be partial differential equations in both time and space variables. These equations, called diffusion equations, will be discussed in Sec. 5.3, In Sec 5.2 we develop Wiener process as the continuous limit of the simple random walk.

It may be noted that there are some measure – theoretic subtleties involved in the passage from the discrete to the continuous case. Their considerations are, however, beyond the scope of this book.

We assume that the process $\{X(t), t \geq 0\}$ is Markovian. Let the cumulative transition probability

be $P(x_0,\ s;\ x,\ t) = \Pr\{X(t) \leq x \mid X(s) = x_0\}, s < 1$ (1.1)

And let the transition probability density p be given by

$P(x_0,\ s;\ x,\ t)dx = \Pr\{x \leq X(t) < x + dx\ Xs{=}x0\}$. (1.2)

For a homogeneous process the transition probability depends only on the length of the interval $(t - s)$ and then the transition probability may be denoted in term of the three parameters, $x_0, x, t - s$.

We denote $\Pr\{x \leq X(t + t_0) < x + dx \mid X(t_0) = x_0\}$ by $p(x_0, x\,; t)dx$

For any t_0, The Chapman – Kolmogorov equation can be written as follows:

$$P(x_0,\ s;\ x,\ t) = \int d_z\, P(x_0,\ s;\ z,\ v)P(z, v\,; x, t).$$

In terms of transition probabilities $P(x_0,\ s;\ x,\ t)$, we have

$$P(x_0,\ s;\ x,\ t) = \int P(x_0,\ s;\ z,\ v)P(z, v\,; x, t)dz.$$

UNIT - VI WIENER PROCESS

6.1. WIENER PROCESS
6.2. DIFFERNTIAL EQUATIONS FOR A WIENER PROCESS
6.3. KOLMOGOROV EQUATIONS

6.1. WIENER PROCESS

Consider that a (Brownian) particle performs a random walk such that in a small interval of time of duration $\Delta t,$ the displacement of the particle to the right or to the left is also of small magnitude $\Delta x,$ the total displacement $X(t)$ of the particle in time t being x. Suppose that the random variable Z_i denotes the length of the i^{th} step taken by the particle in a small interval of time Δt and that

$$\Pr\{Z_i = \Delta x\} = p \; and \; \Pr\{Z_i = -\Delta x\} = q, p + q = 1$$

$0 < p < 1,$ where p is independent of $x \; and \; t.$

Suppose that the interval of length t is divided into n equal subintervals of length Δt and that the displacements $Z_i, i = 1, \dots, n$ in the n steps are mutually independent random variables. Then $n\,.(\Delta t) = t$ and the total displacement $X(t)$ is the sum of $n \; i.i.d.$ random variables $Z_i, \; i.e.$

$$X(t) = \sum_{i=1}^{n(t)} Z_i, \; n \equiv n(t) \frac{t}{\Delta t}.$$

We have $\quad E\{Z_i\} = (p-q)\Delta\, x \; and \; var\;(Z_i) = 4pq\;(\Delta x)^2.$

Hence $\quad E\{X(t)\} = nE(Z_i) = t(p-q)\frac{\Delta x}{\Delta}x, \quad$ (2.1)

And $\quad var\{X(t)\} = n\; var(Z_i) = \frac{4pqr(\Delta x)^2}{\Delta t}.$

To get meaningful result, as $\Delta x \to 0, \Delta t \to 0,$ we must have

$$\frac{(\Delta x)^2}{\Delta t} \to \text{a limit}, \; (p-q) \to \text{a multiple of } \Delta x. \quad (2.2)$$

We may suppose, in particular, that in an interval of length $t, X(t)$ has mean – value function equal to μt and variance function equal to $\sigma^2 t.$ In other words, we suppose that as $\Delta x \to 0, \Delta t \to 0,$ such a way that (2.2) are satisfied, and per unit time

$$E\{X(t)\} \to \mu \text{ and } var\{X(t)\} \to \sigma^2$$
(2.3)

From (2.1) for $t = 1$ and (2.3) we have

$$\frac{(p-q)\Delta x}{\Delta t} \to \mu; \; \frac{4pq(\Delta x)^2}{\Delta t} \to \sigma^2.$$
(2.4)

The relations (2.2) and (2.4) will be satisfied when

$$\Delta x = \sigma(\Delta t)^{1/2},$$
(2.5a)

$$p = \frac{1}{2}\left(1 + \mu(\Delta t)^{1/2}/\sigma\right), \;\; q = \frac{1}{2}\left(1 - \mu(\Delta t)^{1/2}/\sigma\right).$$
(2.5b)

Now since Z_i are i. i. d. random variables, the sum $\sum_{i=1}^{n(t)} Z_i = X(t)$ for large $n(t)(= n)$, is asymptotically normal with mean μt and variance $\sigma^2 t$ (by virtue of the central limit theorem for equal components). Note that here also t represents the length of the interval of time during which the displacement, that takes place is equal to the increment$X(t) - X(0)$.We thus find that for $0 < s < t, \{X(t) - X(s)\}$ is normally distributed with mean $\mu(t - s)$ and variance $\sigma^2(t - s)$. Further, the increments $\{X(s) - X(0)\}$ and $\{X(t) - X(s)\}$ are mutually independent; this implies that $\{X(t)\}$ is a Markov process.

We may now define a **Wiener or a Brownian motion process** as follows:

The stochastic process $\{X(t), t \geq 0\}$ is called a Wiener process (or a Wiener – Einstein process or a Brownian motion process) with drift μ and variance parameter σ^2, if:

(i)$X(t)$ has independent increments, i. e. for every pair of disjoint intervals of time (s, t) and (u, v), where $s \leq t \leq u \leq v$, the random variables $\{X(t) - X(s)\}$ and $\{X(v) - X(u)\}$ are independent.

(ii) Every increment $\{X(t) - X(s)\}$ is normally distributed with mean $\mu(t - s)$ and variance$\sigma^2(t)$.

Note that (i) implies that Wiener process is a Markov process with independent increments and(ii) implies that a Wiener process is Gaussian. Since $\{X(t) - X(0)\}$ is normally distributed with meanμt and variance$\sigma^2 t$, the transition probability density function p of a Wiener process is given by

$$p(x_0, x\, ; t)dx = Pr\{x \leq X(t) x + dx \mid X(0) = x_0\}$$

$$= \frac{1}{\sigma\sqrt{(2\pi t)}} \, exp\left\{-\frac{(x-x_0-\mu t)^2}{2\sigma^2 t}\right\} dx. \qquad (2.6)$$

A Wiener process $\{X(t), t \geq 0\}$ with $X(0) = 0, \mu = 0, \sigma = 1$ is called a *standard* Wiener process.

6.2. DIFFERNTIAL EQUATIONS FOR A WIENER PROCESS

Let $\{X(t), t \geq 0\}$ be a Wiener process. We can consider the displacement in such a process as being caused by the motion of a particle undergoing displacements of small magnitude in a small interval of time. Suppose that $(t - \Delta t, t)$ is an infinitesimal interval of length Δt and that the particle makes in this interval a shift equal to Δx with probability p or a shift equal to $-\Delta x$ with probability $p = 1 - p$. Suppose that $p \, and \, q$ are independent of $x \, and \, t$. Let the transition probability that the particle has a displacement from $x \, to \, x + \Delta x$ at epoch t, given that it started from $x_0$at time0, be $p(x_0, x\,; t)\,\Delta x$. Further suppose that $p(x_0, x\,; t)$ admits of an expansion in Taylor's series, i. e.

$$p(x_0,\; x\; \pm \Delta x\,; t - \Delta t) = p(x_0,\; x\,; t) - \Delta t\,\frac{\partial p}{\partial t} \pm \Delta x \frac{\partial p}{\partial x}$$

$$+\frac{1}{2}(\pm\Delta x)^2 \frac{\partial^2 p}{\partial x^2} + o(\Delta t). \qquad (3.1)$$

From simple probability arguments we have

$$p(x_0,\; x\,; t)\,\Delta x = p\,.\;p(x_0,\; x - \Delta x\,; t - \Delta t)\,\Delta x$$

$$+q\,.p(x_0,\; x + \Delta x\,; t - \Delta t)\,\Delta x. \qquad (3.2)$$

Making use of (3.1), and cancelling out the factor Δx from both sides of (3.2) we get

$$p(x_0,\; x\,; t) = pp(x_0,\; x\,; t) - \;\Delta t\;\frac{\partial p}{\partial t} - \;\Delta x(p - q)\;\frac{\partial p}{\partial x}$$

$$+\frac{1}{2}(\Delta x)^2 \frac{\partial^2 p}{\partial x^2} + o(\Delta t).$$

Divide both sides by Δt. Using (2.4) and (2.5) and taking limits as $\Delta t \to 0, \Delta x \to 0$, we get

$$\frac{\partial}{\partial t} p(x_0,\ x\,;t) = -\mu \frac{\partial}{\partial x} p(x_0,\ x\,;t) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} p(x_0,\ x\,;t. \quad (3.3)$$

This is a partial differential equation in the variables x *and* t, being of first order in t and of the second order in x. The equation is known as the forward diffusion equation of the Wiener process. One can likewise obtain the backward diffusion equation of the process in the form

$$\frac{\partial}{\partial t} p(x_0,\ x\,;t) = -\mu \frac{\partial}{\partial x_0} p(x_0,\ x\,;t) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x_0^2} p(x_0,\ x\,;t. \quad (3.4)$$

The solution of (3.3) (as well as of (3.4)) yields $p(x_0,\ x\,;t)$ as a normal density of the form given in (2.6) (see also Sec. 5.5). It may, however, be easily verified that $p(x_0,\ x\,;t)$ given by (2.6) satisfies (3.3) as well as (3.4). The equation for a Wiener process with drift $\mu = 0$ is known as the *heat equation.*

Note: The partial differential (3.3) [(3.4)] is known as the forward [backward] equation because it involves differentiation in $x[x_0]$. The reason why it is called diffusion equation is given the next section.

It is to be noted that in Sec. 5.2 and 5.3 we have made the following assumptions:

(i) in a small interval of time Δt, the displacement Δx is small (and that $\Delta x = o\left((\Delta t)^{1/2}\right)$);

(ii) $E\{X(t)\} \to \mu t$ in the limit;

(iii) $var\ \{X(t\} \to \sigma^2 t$ in the limit.

The quantity $\mu\ in\ (ii)$ may also interpreted as

$$\lim_{\Delta t \to 0} \frac{E\{X(t+\Delta t) - X(t)\}}{\Delta t} = \mu \quad (3.5)$$

The implies that the infinitesimal mean (i.e. mean overΔt) of the variance of the increment in $X(t)$ exists and is equal to finite quantityσ^2.

For a Wiener process, $\mu\ and\ \sigma^2$ are assumed to be constants, independent of t or of $x(where\ X(t) = x)$. By considering the transition mechanism with $\mu\ and\ \sigma^2$as functions of t or of x or both t and x, we get more general processes for which the equations corresponding to (3.3) and (3.4) will also be more general. We discuss below such equations.

6.3. KOLMOGOROV EQUATIONS

Let $\{X(t), t \geq 0\}$ be a Markov process in continuous time continuous state space. We make the following assumptions: For any$\delta > 0$,

(i) $$Pr\{| X(t) - X(s)| > \delta \mid X(t) = x\} = o(t - s), s < t.$$

In other words, small changes occur during small intervals time.

(ii) $$\lim_{\Delta t \to 0} \frac{E\{X(t+\Delta t) - X(t) \mid X(t) = x\}}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \int_{|y-x| \leq \delta} (y - x) p(x, t; y, t + \Delta t) dy$$

$$= a(t, x).$$

In other words, the limit of the infinitesimal mean of the conditional expectation of the increment of $X(t)$ exists and is equal to $a(t, x)$, which is known as the coefficient.

(iii) $$\lim_{\Delta t \to 0} \frac{E\{[X(t+\Delta t) - X(t)]^2 \mid X(t) = x\}}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \int_{|y-x| \leq \delta} (y - x)^2 \, p(x, t; y, t + \Delta t) dy$$

$$= b(t, x).$$

In other words, the limit of the infinitesimal mean of the variance of the increment of $X(t)$ exists and is equal to $b(t, x)$, which is known as the $diffusion\, coefficient$. A Markov process $\{X(t)\}$ satisfying the above conditions is known as a $diffusion\ equation$. We give below the equations .

Let $\{X(t), t \geq 0\}$be a Markov process satisfying $(i), (ii) and (iii.)$ If its transition p.d. f.

$p(x_0,\ t_0;\ x,\ t)$ possesses continuous partial derivatives

$$\frac{\partial p}{\partial t}, \frac{\partial p}{\partial x}(a(t, x)p), \frac{\partial^2}{\partial x^2}(b(t, s)p),$$

Then $p(x_0,\ t_0;\ x,\ t)$ satisfies the forward Kolmogorov equation

$$\frac{\partial p}{\partial t} = -\frac{\partial p}{\partial x}(a(t, x)p) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(b(t, x)p). \qquad (4.1)$$

This equation is also known as the$Fokker - Planck\ equation$. Suppose that $p(x_0,\ t_0;\ x,\ t)$ possesses continuous partial derivatives

$$\frac{\partial p}{\partial t_0}, \frac{\partial p}{\partial x_0}, \frac{\partial^2 p}{\partial x_0^2};$$

Then $p(x_0,\ t_0;\ x,\ t)$ also satisfies the backward Kolmogorov equation

$$\frac{\partial p}{\partial t_0} = -a(t_0, x_0)\frac{\partial p}{\partial x_0} - \frac{1}{2}\ b(t_0, x_0)\frac{\partial^2 p}{\partial x_0^2}. \quad (4.2)$$

The diffusion equations for $X(t)$ were first derived by Kolmogorov. Feller showed that under suitable restrictions the equations admit of a unique solution. Fortet established some very interesting and important properties of the solutions.

Particular case: If the process is homogeneous, then

$$p(x_0,\ t_0\,; t, x) = p(x_0, x\ ;\ t - \ t_0),$$

and$a(t,x), b(t,x)$ are independent of t.

If the process is additive, i. e. given that$X(\ t_0\ ) =\ x_0$, the increment $\{X(t), X(\ t_0\ )\}$, depends only on $t_0\ \ and\ t$ (and not on x_0), then

$$p(x_0,\ t_0\,; t, x) = p(x - x_0\ ;\ t_0,\ t)\ and\ a(t,x),$$ $b(t,x)$

Are independent of x.

The Kolmogorov equations, in these case, can be easily deduced from the general equations.

Unit – VII FIRST PASSAGE TIME DISTRIBUTION FOR WIENER PROCESS

7.1. FIRST PASSAGE TIME DISTRIBUTION FOR WIENER PROCESS

The possible realizations of a stochastic process are called *sample paths or trajectories.* The structure and the properties of the same paths of a Brownian motion or Wiener process are the subject matter of deep study. Without entering into the subtleties(which are beyond the scope of this work), we discuss here some results of Wiener process, using the property that the sample paths are continuous functions. We also make use of the simple but powerful ‘ reflection principle’. The principle relates to the fact that there is a one –to-one correspondence between all paths from $A(a_1,\ a_2) to\ B(b_1,\ b_2)$ which touch or cross the x-axis and all paths from $A'(a_1,\ -a_2)\ to\ B$(see, Feller, Vol I for details). We shall first consider the following from which the distribution of the first passage time will be derived.

7.1.1 Distribution of the Maximum of a Wiener Process

Lemma: Let $\{X(t), 0 \le t \le T\}$ be a Wiener process with $X(0) = 0\ and\ \mu = 0. Let\ M(T)$ be the maximum of $X(t) in\ 0 \le t \le T, i.e. M(T) = max_{0\le t\le T}\ X(t)$. Then for any $a > 0$

$$\Pr\{M(T) \ge a\} = 2\Pr\{X(T\} \ge a\}.$$

(This result was first obtained by Bachelier(1900).)

Proof: Consider the collection of sample path$X(t), 0 \le t \le T$ such that $X(T\} \ge a$. Since $X(0) = 0\ and\ X(t)$ is continuous, there exists a time τ at which $X(t)$ first attains the value a (or $X(t)$this first the value a). The time T_a is itself a random variable. For $> T_a$, $X_a(t)$ gives below

$$X_a(\text{t}) = \begin{cases} X(t), & t < T_a \\ 2a - X(t), & t > T_a \end{cases}$$

gives reflection of $X(t)$ about the line$x = a$. Note that $X_a(\mathrm{T}) \leq a, and\ that\ M(T) = max_{0\leq t\leq T} X(T) > a$ and $M_a(\mathrm{T}) = max_{0\leq t\leq T} X_a(t) \geq a$; further, by symmetry the sample paths $X(t)\ and\ X_a(\mathrm{t})$ have the same probability of occurrence. From reflection principle, it follows that corresponding to every sample path $X(t)$ for which $X(T) \geq a$, there exist two sample paths such that $M(T) \geq a$, Further, its converse is also true, viz., every sample paths $x(t)$ for which $M(T) \geq a$ corresponds to two sample paths $X(t)$ with equal probability, one of the paths being such that $X(T) > a$, unless $\{X(T) = a\}$, whose probability is zero. In fact, the set $\{M(T) = a\}$ is the union of three disjoint sets

$$\{M(T) \geq a, X(t) > a\},$$

$$\{M(T) \geq a, X(t) < a\},$$

and $$\{M(T) \geq a, X(t) = a\},$$

The probability of the third set is zero, while the two are mapped onto one another by reflection about the line $x = a$ after the time T_a. Thus we have

$$\Pr\{M(T) \geq a,\} = 2\Pr\{M(T) \geq a\}.$$

The above gives a heuristic proof of the lemma; as already indicated, a rigorous proof involves considerations beyond the scope of this book (see Karlin and Taylor, Iosifescu and Tautu).

Let $\{X(t), t \geq 0\}$ be a Wiener process with $X(0) = 0, \mu = 0\ and\ M(t) = max_{0\leq s\leq t} X(s)$. Then from the lemma, we get, for $t > 0, \sigma = 1$.

$$\Pr\{M(t) \geq a,\} = 2\Pr\{M(t) \geq a\}$$

$$= \frac{2}{\sqrt{(2\pi t)}} \int_a^\infty exp(-x^2/2t)dx$$

(5.1a)

$$= \frac{2}{\sqrt{(2\pi)}} \int_{a/\sqrt{t}}^\infty exp(-y^2/2)dy$$

$$\left(by\ changing\ the\ variable\ to\ y = {}^{x}\!/\!_{\sqrt{t}}\right)$$

$$= 2\left\{1 - \frac{2}{\sqrt{(2\pi)}} \int_{-\infty}^{a/\sqrt{t}} exp(-y^2/2)dy\right\}$$

$$= 2\{1 - \Phi(a/\sqrt{t})\}.$$

Φ being the distribution function of the standard normal variate.

By changing the variable to $s = \frac{a^2 t}{x^2}$, (5.1a) can be written as

$$\Pr\{M(t) \geq a\} = \frac{a}{\sqrt{(2\pi)}} \int_0^t s^{-3/2} \, exp\left(-\frac{a^2}{2s}\right) ds, s > 0.$$

(5.1d)

7.1.2**Distribution of the First Passage Time to a Fixed Point**

We can use the lemma to obtain the distribution of the random variable T_a, the first passage time to a fixed point $a(> 0)$ (or the time of hitting a fixed point a first), for a Wiener process$\{X(t)\}\, with\, X(0) = 0, \mu = 0$. The time $T_a\, for\, X(t)$ to hit the level a first will be less than $t\, iff\, M(t) = max_{0 \leq s \leq t}\, X(s)$ in that time is at least a.

Thus for $t > 0$ $$Pr\{M(t) \geq a\} = Pr(T_a \leq t).$$ (5.2)

Hence the distribution function $F(t) = \Pr\{T_a \leq t\}$ is

$$F(t) = \frac{2}{\sqrt{(2\pi t)}} \int_a^\infty exp(-x^2/2t) dx$$ (5.3a)

$$= 2\left\{1 - \Phi(a/\sqrt{t})\right\}$$ (5.3b)

$$= \frac{a}{\sqrt{(2\pi)}} \int_0^t s^{-3/2} \, exp\left(-\frac{a^2}{2s}\right) ds, s > 0.$$ (5.3c)

The density function of T_a is obtained by differentiating (5.3) with respect to t. Differentiation of (5.3c) (and also 5. 3b) readily gives

$$f_{T_a}(t) = F'(t) = \frac{a}{\sqrt{(2\pi)}} t^{-3/2} exp\left(-\frac{a^2}{2t}\right), \; t > 0$$ (5.4)

It may be easily verified that the Laplace transform is

$$\overline{f}(s) = \int_0^\infty e^{-st} f(t) dt$$

$$= \exp\left\{-a\sqrt{(2s)}\right\}.$$

It can be seen that on moment of T_a exists finitely.

Let us find the density function of $M(t)$. The distribution function is

$$G(a) = \Pr\{M(t) \le a\}$$

$$= 1 - \Pr\{M(t) > a\}$$

$$= 1 - 2\Pr\{X(t) \ge a\}$$

$$= 1 - \frac{2}{\sqrt{(2\pi t)}} \int_a^\infty exp\,(-x^2/2t)\,dx$$

(5.5a)

$$= 1 - 2\{1 - \Phi(a/\sqrt{t})\}$$

(5.5b)

$$= 2\Phi(a/\sqrt{t}) - 1.$$

Differentiating (5.5a)(or 5.5b) with respect to a, we get the density function

$$g_M(a) = G'(a) = \frac{2}{\sqrt{(2\pi t)}}\, exp\left(-\frac{a^2}{2t}\right),\ a > 0.$$

(5.6)

The result given above (for $\sigma = 1$) can be suitably modified for any $\sigma > 0$.

Note 1: We have obtained the distribution of T_a by using the lemma which gives a relation between the distributions of the Wiener process $X(t)$ and its maximum $M(t)$. However, the distribution of T_a can be obtained directly without bringing in the distribution of the maximum. IN fact, it can be directly shown that, if $X(0) = 0, a > 0$, then

$$\Pr\{T_a \le t\} = 2\Pr\{X(t\} \ge a\}.$$

(5.7)

For a proof of the above, see Prohorov and Rozanov, who use conditional expectations to obtained by using (5.2).

Note 2: For an alternative approach to the distribution of $T_a$using differential equations,see Cox and Miller, who obtain the distribution for any $\mu \ge 0\ and\ \sigma > 0$. The density function $f_{T_a}(t)$ $(for\ X(0) = 0, \mu \ge 0, \sigma > 0)$ of T_a is found to be

$$f_{T_a}(t) = \frac{a}{a\sqrt{(2\pi x^3)}} \exp\left\{-\frac{(a-\mu x)^2}{2\sigma^2 x}\right\}, x > 0;$$

(5.8)

And its Laplace transform is

$$\bar{f}(s) = exp\left[(a/\sigma^2)\left\{+\mu - \sqrt{(\mu^2 + 2s\sigma^2)}\right\}\right].$$

The mean and the variance for T_a for $\mu \neq 0$ are given by

$$E\{T_a\} = a/\mu \ and \ var\{T_a\} = a\sigma^2/\mu^3.$$

Note 3:

The function(5.8) with $\mu > 0$ is the density function of the distribution of the passage time of Brownian motion with a positive drift. This distribution having density function (5.8) is known as *inverse Gaussian*distribution because of the inverse relationship between the cumulant generating function of this distribution and that of normal distribution. Such a distribution was also obtained by Wald as the limiting form for the distribution of the sample size in a sequential probability ratio test. For properties of this distribution, see Johnson and Kotz (1970), and for statistical applications, see Folks and Chhikara (1978).

Example 5(a).

Suppose that $\{X(t), 0 < t,\}$ is a Wiener process with $X(0) = 0, and\ \mu = 0$. Then

$$Pr\{X(t) \leq x\} = Pr\{X(t)/\sigma\sqrt{t} \ \leq x/\sigma\sqrt{t}\} = \Phi(x/\sigma\sqrt{t}).$$

Consider the process $Y(t) = tX(1/t)\ in\ 0 < t \leq 1\ with\ Y(0) = 0.$

We have $E\{Y(t) = 0\ and\ var\{Y(t)\} = t^2(\sigma^2/t) = \sigma^2/t.$

Further, $Pr\{Y(t) \leq y\} = Pr\{tX(1/t) \leq y\} = Pr\left\{\frac{X(1/t)}{\aleph\sqrt{(1/t)}} \leq \frac{y(t)}{\aleph\sqrt{(1/t)}}\right\}$

Thus $\{Y(t), 0 < t \leq 1\}\ with\ Y(0) = 0$ is also a Wiener process with $\mu = 0$ and variance $\sigma^2\ t$.

Example 5(b).

Consider a Wiener process $\{X(t)\}$ with $X(0) = 0$. Its mean value function is μt and variance function $\sigma^2\ t$. For $0 < s < t$, the covariance function is

$$C(s,t) = cov\{X(s), X(t)\} = cov\{X(s), X(s) + X(t) - X(s)\}$$

$$= cov\{X(s), X(s)\} + cov\{X(s), X(t) - X(s)\}$$

$$= cov\{X(s), X(s)\},$$

Since the process has independent increments. Thus

$$C(s,t) = var\{\,X(s)\} = \sigma^2 s$$

The process is not covariance stationary even when $\mu = 0$.

Example 5(c).

Suppose that $\{X(t), t > 0\}$ is a Wiener process with $X(0) = 0$. Its first passage time, $T_a \; to \; a$ has same distribution as $1/\mu^2$, where u is a normal variate with mean 0 and s. d. σ/a. For, the distribution function of $1/\mu^2$ for $t > 0$, is

$$F(t) = Pr\{1/\mu^2 \le t\} = Pr\{\mu^2/(\sigma/a)^2 \ge a^2/\sigma^2 t\}$$

$$= Pr\{(au/\sigma) \ge (a/\sigma\sqrt{t})\} + Pr\{(au/\sigma) \le$$

$-a\sigma t$

$$= 1 - \Phi(a/\sigma\sqrt{t}) + \Phi(-a/\sigma\sqrt{t})$$

$$2\left(1 - \Phi(a/\sigma\sqrt{t})\right).$$

Which is the distribution function of T_a (see equation (5.3b)). The distribution of $1/\mu^2$ is *inverse Gaussian.*

Example 5(d).

Let $\{X(t), t \ge 0\}$ be a Wiener process with $\mu = 0 \; and \; X(0) = 0$. To find the distribution of T_{a+b} for $0 < a < a + b$.

Suppose that t_a is a value of $T_a, i.e. X(t)$ reaches the level a for the first time a epoch t_a, We may then consider that the process starts at $(t_a,\; a)$ and reaches the level $(a + b)$, which is b units higher than a. Suppose that $T_{a+b} - T_a$ is the duration of the interval at the ends of which $X(t)$ first reaches the level a and then reaches first the level $a + b$. Then $T_a \; and \; (T_{a+b} - T_a)$ are independent random variables denoting first passage times to $a \; and \; b$ respectively. The L.T. of the p. d. f. of T_a is

$$\overline{h_a}(s) = exp\left\{-\sqrt{(2s)}/\sigma\right\}$$

and that of T_a is $$\overline{h_b}(s) = exp\left\{-b\sqrt{(2s)}/\sigma\right\}.$$

Thus the L. T. of the p. d. f. of T_{a+b} is

$$\overline{h_a}(s)\overline{h_b}(s) = exp\left\{-(a+b)\sqrt{(2s)}/\sigma\right\} =$$
$\overline{h_{a+b}}(s)$.

7.2 ORNSTEIN – UHLENBECK PROCESS

We have seen that for a Wiener process $\{X(t)\}$, the displacement Δx, in a small interval of time Δt is also small, being of $[O(\sqrt{\Delta t})]$. The velocity which is of $O(\sqrt{\Delta t}/\Delta t) = O(1/\sqrt{\Delta t})$ tends to infinity as $\Delta t \to 0$. Thus the Wiener process does not provide a satisfactory model for Brownian motion for small values of t, although for moderate and large value of t it does so. An alternative model which holds for small t proposed by Ornstein and Uhlenbeck in 1930. Here instead of the displacement

The equation of motion of a Brownian particle can written as

$$dU(t) = -\beta U(t)dt + dF(t), \qquad (6.1)$$

Where $-\beta U(t)$ represents the systematic part due to the resistance of the medium and $dF(t)$ represents the random component. It is assumed that these two parts are independent and that $F(t)$ is a Wiener process with drift $\mu = 0$ and variance parameter σ^2. The Markov process $\{U(t), t \geq 0\}$ is such that in a small interval of time the change in $U(t)$ is also small. Since $F(t)$ is a Wiener process, we have from (6.1)

$$lim_{\Delta t\to 0}\frac{E\{U(t+\Delta t)-U(t)|\ U(t)=u\}}{\Delta t}$$

$$= -\beta u + lim_{\Delta t\to 0}\frac{E\{\Delta F(t)\}}{\Delta t}$$

$$= -\beta u,$$

and

$$lim_{\Delta t\to 0}\frac{var\{U(t+\Delta t)-U(t)|\ U(t)=u\}}{\Delta t}$$

$$= lim_{\Delta t\to 0}\frac{(\Delta t)^2}{\Delta t} + \frac{var\{\Delta f(t)\}}{\Delta t},$$

$$= \sigma^2.$$

In other words, the limits exist. So the process $\{U(t), t \geq 0\}$ is a diffusion process and its transition p. d. f. $p(u_0; u, t)$ satisfies the forward Kolmogorov equation (4.1) with $a(u,t) = -\beta u$ and $b(u,t) = \sigma^2$.

This is, p satisfies the differential equation

$$\frac{\partial p}{\partial t} = \beta \frac{\partial}{\partial u}(up) + \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial u^2}. \qquad (6.2)$$

Let us assume that $U(0) = u$ and that as $u_0 \to \pm\infty,\ p \to 0$ and $\left(\frac{\partial p}{\partial u}\right) \to 0$. The solution of (6.2) gives p, the transition p. d. f. of $U(t)$. It is more convenient to consider the equation corresponding to (6.2) in the characteristic function of $p, i.e.$

$$\Phi(u_0;\ \theta, t) = \int_{-\infty}^{\infty} e^{i\theta u}\, p(u_0; u,)du.$$

We have $\int_{-\infty}^{\infty} e^{i\theta u} \frac{\partial}{\partial u}(up)du = e^{i\theta u}\, up\, |_{-\infty}^{\infty} - \int_{-\infty}^{\infty} i\theta\, e^{i\theta u}\, u\, p\, du$

$$= -\theta \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} e^{i\theta u}\, p\, du$$

$$= -\theta \frac{\partial \phi}{\partial \theta};$$

$$\int_{-\infty}^{\infty} e^{i\theta u} \frac{\partial^2 p}{\partial u^2} = e^{i\theta u} \frac{\partial p}{\partial u} |_{-\infty}^{\infty} - \ i\theta\ \int_{-\infty}^{\infty} e^{i\theta u} \frac{\partial p}{\partial u} du$$

$$= -\, i\theta \left\{ e^{i\theta u}\, p\, |_{-\infty}^{\infty} - \ i\theta \int_{-\infty}^{\infty} e^{i\theta u}\, p\, du \right\}$$

$$= -\theta^2 \phi.$$

The equation (6.2) then becomes

$$\frac{\partial \phi}{\partial t} + \beta\theta \frac{\partial \phi}{\partial \theta} = -\frac{1}{2}\sigma^2\theta^2\phi. \qquad (6.3)$$

The equation (6.3) is of Lagrange type. It can be shown that

$$\phi(u_0; \theta, t) = exp\left\{ i\theta u_0 e^{-\beta t} - \frac{1}{4\beta}\theta^2\sigma^2 (1 - e^{-2\beta t}\right. \qquad (6.4)$$

This is the characteristic function of normal distribution with $m(t) = u_0 e^{-\beta t}$ and variance function $\sigma^2(t) = \sigma^2\left(1 - e^{-2\beta t}\right)/4\beta$.

In other words, the transition p. d. f. p is normal with mean value function $m(t)$ and variance function $\sigma^2(t)$ and p can be written as:

$$p(u_0; u, t) = \frac{1}{\sqrt{(2\pi\sigma^2(t))}}\, exp\{-(x - m(t))^2/2\sigma^2(t)\}.$$

Thus the process $\{U(t), t \geq 0\}$ is a Gaussian process with mean value function $m(t)$ and variance function $\sigma^2(t)$. $\{U(t), t \geq 0\}$ is a Markov process but it does not possess independent increments like the Wiener process. $\{U(t), t \geq 0$ is known as *Ornstein – Uhlenbeck process (O - U. P.).* For large t, $mt \to 0$ and $\sigma^2(t) \to \frac{\sigma^2}{2\beta}$, *i.e.* the distribution of velocity is normal with mean 0 and variance $\frac{\sigma^2}{2\beta}$. We thus get an equilibrium distribution and $U(t)$ is said to in statistical equilibrium. For small t, $m(t) \to u_0$ and $\sigma^2(t) \to \sigma^2 t$.

Example 6(a).

Joint distribution of $U(t)$ *and* $U(t+\tau)$ when $U(t)$ is in equilibrium. For large t, the limiting distribution of $U(t)$ is normal with mean o and variance $\frac{\sigma^2}{2\beta} = \sigma_0^2$. The conditional distribution of $U(t+\tau)$, given $U(t) = u$, is normal with mean $ue^{-\beta\tau}$ and variance $\sigma^2\tau = \sigma^2\left(1 - e^{-2\beta t}\right)/2\beta$. Thus the unconditional distribution of $U(t+\tau)$ has the following density

$$h(x) = \frac{1}{\sqrt{(2\pi\sigma_0^2)}} \int_{-\infty}^{\infty} \left[exp\left(-\frac{x_0^2}{2\sigma_0^2}\right) \frac{1}{\sqrt{(2\pi\sigma^2(t))}} \times\right.$$

$exp-12\sigma2(t)x-x0e-\beta t2\ dx0$

$$= \frac{1}{\sqrt{(2\pi\sigma_0^2)}}\ exp\left(-\frac{x^2}{2\sigma_0^2}\right).$$

Thus the unconditional distribution of $U(t+\tau)$ is Gaussian, has mean 0 and variance $\sigma_0^2$and the unconditional distribution of $U(t+\tau)$ is the same as the equilibrium distribution of $U(t)$. The joint distribution of $U(t)$ *and* $U(t+\tau)$has the density

$$F(x,y) = \frac{1}{\sqrt{(2\pi\sigma_0^2)}}\ exp\left(-\frac{x^2}{2\sigma_0^2}\right) \cdot \frac{1}{\sqrt{(2\pi\sigma^2(t))}}\ exp\left\{-\frac{1}{2\sigma^2(t)}\left(y - \right.\right.$$

$xe-\beta t2$

$$\frac{1}{2\pi\sigma_0^2\sqrt{(1-c^2)}}\ exp\left\{-\frac{1}{2\sigma_0^2(1-c^2)}\left(y^2 - 2axy + \right.\right.$$

$x2,$

Where $c = e^{-\beta t}$.

It follows that $U(t)$ *and* $U(t+\tau)$ have a bivariate Gaussian distribution with

$$E\{U(t)\} = E\{U(t+\tau)\} = 0,$$

$$cov\{U(t), U(t+\tau)\} = \sigma_0^2 e^{-\beta\,|\,\tau\,|}$$

and $$var\{U(t)\} = \sigma_0^2.$$

The mean and the variance of $U(t)$ are finite and the covariance function $cov\{U(t), U(t+\tau)\}$ is a function of the absolute difference only. Hence $\{U(t)\}$ is covariance stationary. Again, since $\{U(t)\}$ is Gaussian, $\{U(t)\}$ is strictly stationary.

Note that Wiener process is not covariance stationary.

Example 6(b). *The O – U. P. as a transformation of a Wiener process:*

Let $\{X(t), t \geq 0\}$, be a standard Wiener process. Let

$$Y(t) = \frac{1}{\sqrt{g}\,(t)}\, X\big(ag(t)\big), a > 0.$$

And let the (non - random) function $g(t)$ be positive, strictly increasing with $g(0) = 1$.

We have $$E\{Y(t)\} = 0$$

$$var\{Y(t)\} = \frac{1}{g(t)}\, var\{X(ag(t))\}$$

$$= \frac{1}{g(t)}\{ag(t)\} = a.$$

Since $(X(t_1), \dots, X(t_n))$ is multivariate normal, so also is $\big(X\big(a\,g(t_1)\big)\,, \dots, X\big(a\,g(t_n)\big), and\ for\ \tau > 0,$

$$cov\{Y(t),\ Y(t+\tau)\} = cov\left\{\frac{X\big(a\,g(t)\big),\ X(a\,g(t+\tau))}{(g(t)g(t+\tau))^{1/2}}\right\}$$

Since, for the Wiener process $\{X(t)\}, cov\,\{X(t+\tau)\} = varX(t)$ (see Example 5(b)).

Now $$var\{X(ag(t))\} = var\big\{\sqrt{g}\,(t)Y(t)\big\}$$

$$= g(t)var\,\{Y(t)\} = ag(t)$$

and thus $$cov\{Y(t),\ Y(t+\tau)\} = a\{g(t)/g(t+\tau)\}^{1/2}$$

and for $t = 0, the\ covariance\ equals\ a/\{g(\tau)\}^{1/2}$.

The process $\{Y(t)\}$, which has finite mean and variance, will be covariance provided $cov\{Y(t),\ Y(t+\tau)\}$ depends only on τ. Thus we must have

$g(t+\tau) = g(\tau).g(t)$, and in order to satisfy this equation, $g(t)$ must be an exponential function, say, $g(t) = e^{2\beta t}(\beta > 0)$. We have then

$$cov\{Y(t),\ Y(t+\tau)\} = ae^{-\beta t}(\tau > 0),$$

and thus we find that for $a > 0,\ \beta > 0$

$$Y(t) = e^{-2\beta t}X(ae^{2\beta t})$$

Is a stationary Gaussian Markov process. In other words, $Y(t)$ has the structure of an Ornstein – Uhlenbeck process.

7.2.1 Remarks:

The concepts of O – U, process have been applied extensively in Finance, Economics and Management. O – U. process has been used as models for continuous control system, for buffer stock control, for continuous industrial processes in chemical plants, for process control in thermal plants and so on in industrial management; for pricing in a large system of cash bonds and so on in financial measurement; as well as for interest rate behavior and so on in economics. For application of diffusion processes in Finance and Economics, refer to Mallaris and Brock (1982).

Problems:

5.1 If $X(t)$, with $X(0)\ and\ \mu = 0$, is a Wiener process, show that $Y(t) = \sigma X(t/\sigma^2)$ is also a Wiener process. Find its covariance function.

5.2 If $X(t)$, with $X(0)\ and\ \mu = 0$, is a Wiener process and $0 < s < t$, show that for at least one τ satisfying $s \leq \tau \leq t$,

$$\Pr\{X(\tau) = 0\} = \left(\frac{2}{\pi}\right) cos^{-1}\left((s/t)^{1/2}\right).$$

5.3 Let $X(t)$, with $X(0) = 0$, be a standard Wiener process and let T_a be the first passage time of $X(t)$. Show that $T_a\ and\ a^2T_1$ are identically distributed. If $Z_i, i = 1, 2, \dots, n$ are i. i. d. as T_1 then show that $\frac{\sum Z_i}{n}\ and\ Z_i$ are identically distributed.

UNIT – VIII BRANCHING PROCESS

8.1 BRANCHING PROCESS INTRODUCTION

The history of the study of branching processes dates back to 1874, when a mathematical model was formulated by Galton and Watson for the problem of ‘ extinction of families’. The model not attract much attention for a long time; the situation gradually changed and during the last 60 years much attention has been devoted to it. This is because of the development of interest in the applications of probability theory, in general, and also because of the possibility of using the models in a variety of biological, physical and other problems where one is concerned with objects that can generate objects of similar kind; such objects may be biological entities, such as human beings, animals, genes, bacteria and so on, which yield new neutrons under a nuclear chain reaction or in the process of nuclear fission.

We consider first the discrete time case. Suppose that we start with an initial set of objects(or individuals) which form the 0^{th} generation – these objects are called *ancestors.* The off-springs reproduced or the objects generated by the objects of the 0^{th} generation are the ‘ direct descendants’ of the ancestors, and are said to form the 1^{st} generation; the objects generated by these of the 1^{st} generation (or the direct descendants of the 1^{st}generation) from the 2^{nd} generation, and so on, the direct descendants of the r^{th} generation form the $(r+1)st$ generation. The number of objects of the r^{th} generation $(r = 0, 1, 2, ..)$ is a random variable. We assume that the objects reproduce independently of other objects, i. e., there is no interference.

8.1.1 Definition.

Let the random variables$X_0, X_1, X_2, ..$ denote the sizes of (or the numbers of objects in) the $0^{th}, 1^{st}, 2^{nd}, ..,$ generations respectively. Let the probability that an object (irrespective of the generation to which it belongs) generates k similar objects be denoted by $p_k, where\ p_k \geq 0, k = 0, 1, 2, .., \sum_k p_k = 1.$

The sequence $\{X_n, n = 0, 1, 2, .., \}$ constitutes $a\ Galton - Watson\ branching\ process$ (or simply a G. W. branching process) with off – spring distribution $\{p_k\}$.

The process is also called *Bienayame – Galton – Watson process*, in recognisation of an even earlier work by Bienayame.

Our interest lies mainly in the probability distribution of X_n and the probability that $X_n \to 0$ for some $n, i.e.$, the probability of ultimate extinction of the family.

8.1.2 Note 1:

Unless otherwise stated, we shall assume that $X_0 = 1, i.e.$, the process starts with a single ancestor.

8.1.3 Note 2:

The sequence (X_n) forms a Markov chain with transition probabilities

$$p_{ij} = Pr\{X_{n+1} = j \mid X_n = i\}, i, j = 0, 1, 2, ..$$

It is however not always easy to specify p_{ij}.

8.1.4 Note 3:

The generating functions are very useful in the study of branching processes.

8.2 PROPERTIES OF GENERATING FUNCTIONS OF BRANCHING PROCESSES

Another definition given is as follows:

A Galton – Watson process is a Markov chain $\{X_n, n = 0, 1, 2, ..\}$ having state space N(set of non – negative integers), such that

$$X_{n+1} = \sum_{r=1}^{X_n} \varsigma_r, \quad (2.1)$$

Where ς_r are i. i. d. random variables with distribution $\{p_k\}$.

Let $$P(s) = \sum_k Pr\{\varsigma_r = k\} s^k = \sum_k p_k s^k \quad (2.2)$$

be the p. g. f. of $\{\varsigma_r\}$ and let

$$P_n(s) = \sum_k Pr\{X_n = k\}s^k, n = 0, 1, 2, ..$$

(2.3)

be the p. g. f. of $\{X_n\}$.

We assume that $X_0 = 1$; clearly $P_0(s) = s$ and $P_1(s) = P(s)$. The r. v.' a X_1 and ς_r both give off – spring distribution.

Theorem 8.2.1

We have $$P_n(s) = P_{n-1}(P(s))$$ (2.4)

and $$P_n(s) = P(P_{n-1}(s)).$$ (2.5)

Proof:

We have, for $n = 1, 2, ...$

$$Pr\{X_n = k\} = \sum_{j=0}^{\infty} Pr\{X_n = k \mid X_{n-1} = j\}. Pr\{X_{n-1} = j$$

$$= \sum_{j=0}^{\infty} Pr\left\{\sum_{r=1}^{j} \varsigma_r = k\right\}. Pr\{X_{n-1} = j\}$$

So that, $$P_n(s) = \sum_{k=0}^{\infty} Pr\{X_n = k\}\, s^k$$

$$= \sum_{k=0}^{\infty} s^k \left[\sum_{j=0}^{\infty} Pr\left\{\sum_{r=1}^{i} \varsigma_r = k\right\} Pr\{X_{n-1} = j$$

$$= \sum_{j=0}^{\infty} Pr\{X_{n-1} = j\}\left[\sum_{r=1}^{i} Pr\{\varsigma_1 + \varsigma_2 + \cdots + \varsigma j = k s k.$$

The expression within square brackets, being the p. g. f. of the sum $\varsigma_1 + \cdots + \varsigma_j$ of j i. i. d. random variables each with p. g. f $P(s)$, equals $[P(s)]^j$. Thus

$$P_n(s) = \sum_{j=0}^{\infty} Pr\{X_{n-1} = j\}\, [P(s)]^j$$

$$= P_{n-1}(P(s)).$$

Thus we get (2.4). Putting $n = 2, 3, 4, ...$, we get, when $X_0 = 1$,

$$P_2(s) = P_1\big(P(s)\big) = P\big(P(s)\big),\ P_3(s) = P_2(P(s)), P_4(s) = P_3(P(s))$$

and so on, Iterating (2.4) we get

$$P_n(s) = P_{n-1}(P(s)) = P_{n-2}(P(P(s)))$$
(2.6)

$$= P_{n-2}(P_2(s)).$$

For $n = 3$, $P_3(s) = P_1(P_3(s)) = P(P_3(s))$.

Again iterating (2.6), we get

$$P_n(s) = P_{n-3}(P(P_3(s))) = P_{n-3}(P_3(s)),$$

and for $n = 4$, $P_4(s) = P_1(P_3(s)) = P(P_3(s))$.

Thus $P_n(s) = P_{n-k}(P_k(s)), k = 0, 1, 2, .., n$,

and for $k = n - 1$ $P_n(s) = P_1(P_{n-1}(s)) = P(P_{n-1}(s))$.

Thus we get (2.5).

Note:that even when$X_0 = i \neq 1$, the relation (2.5) holds but (2.4) does not hold.

8. 2.2 Remarks.

Theorem 8.2.1 , Could be used to find the **moments of X_n**.

We have

$$P'(1) = E(\varsigma_r) = E(X_1) = m(say).$$

Theorem 8.2.3

If $m = E(X_1) = \sum_{k=0}^{\infty} kp_k,\ and\ \sigma^2 = var(X_1)$ then

$$E\{X_n\} = m^n$$
(2.7)

and $$var(X_n) = \frac{m^{n-1}(m^n-1)}{m-1}\sigma^2, \qquad if\ m \neq 1$$
(2.8)

$$= n\sigma^2, \qquad if\ m = 1.$$

Proof: Differentiating (2.4,) we get

$$P_n'(s) = P_{n-1}'(P(s))P'(s)$$

whence $P_n'(1) = P_{n-1}'(1)P'(1) = mP_{n-1}'(1)$

and on iterating $P_n'(1) = m^2 P_{n-2}'(1)$

$$m^{n-1} P'(s) = m^n.$$

Thus $E(X_n) = P_n'(1) = m^n.$

Differentiating (2.5) twice and proceeding in a similar fashion, one can find the second moment $P_n''(1)$, and thus the variance of X_n in the form (2.8).

One can likewise proceed to get higher moments of X_n.

Alternatively, the mean and the variance of X_n can be obtained by nothing that X_{n+1} is the sum of a random number of i. i. d. random variables, and using standard formulas.

Let $m \neq 1$. We can use the Corollary to Theorem 1.3 to find $E(X_{n+1})$.

Since $X_{n+1} = \sum_{r=1}^{X_n} \varsigma_r$,we have

$$E(X_{n+1})\, E(\varsigma_r)\, E(X_n) = mE(X_n).$$

The solution of the difference equation is given by

$$E(X_n) = Cm^n, \qquad n = 1, 2, 3, \dots$$

Since $E(X_1) = E(\varsigma_r) = m, C = 1. Thus\ E(X_n) = m^n.$

Using the given in (relation (1.20)), chapter 1, we get

$$var(X_{n+1}) = E(X_n)\, var(\varsigma_r) + [E(\varsigma_r)^2]\, var(X_n)$$
(2.9a)

$$= m^n \sigma^2 + m^2 var(X_n)$$
(2.9b)

We can find $var(X_n)$ from (2.9b) either by induction or by solving the non – homogeneous difference equation. We employ the latter method. A particular solution of the difference equation (2.9b) is given by $var(X_n) = \frac{\sigma^2 m^n}{m - m^2}$

And a general solution of the homogeneous equation corresponding to (2.9b) is given by $var(X_n) = A(m^2)^n$, where A is a constant. Thus complete solution of (2.9b) is given by

$$var(X_n) = A(m^2)^n + \frac{\sigma^2 m^n}{m - m^2}$$
(2.10)

Nothing that $var(X_1) = \sigma^2$, we get $A = \sigma^2/\{m(m-1)\}$ so that (2.10) yields

$$var(X_n) = \frac{m^{n-1}(m^n-1)}{m-1}\sigma^2, n = 1, 2, \ldots$$

The result holds for all n and $m \neq 1$. By taking limit as $m \to 1$, one gets the corresponding result for $m = 1$.

When $m = 1$, then $var(X_n) = n\sigma^2$.

When $m = 1$, the variance of X_n increases linearly and when $m > 1$ $(m < 1)$ it increases geometrically with n.

According as $m < 1, = 1, or > 1$, the Galton – Watson process is referred to as *subcritical, critical orsupercritical respectively*. □

We now come to the problem originally posed by Galton.

BLOCK III

UNIT – IX PROBABILITY OF EXTINCTION AND STOCHASTIC IN M/M/1 – MODEL

9.1 PROBABILITY OF EXTINCTION

9.1.1 Definition:

By extinction of the process it is meant that the random sequence $\{X_n\}$ consists of zeros for all except a finite number of values of n. In order words, extinction occurs when $Pr\{X_n = 0\} = 1$, for some value of n. Clearly, if $X_n = 0 \; for \; n > m; also \; Pr\{X_{n+1} = 0 \mid X_n = 0\} = 1$.

Theorem 9.1.2

If $m \leq 1$, the probability of ultimate extinctions is 1. If $m > 1$,the probability of ultimate extinction is positive root less than unity of the equation

$$P(s) = s. \tag{3.1}$$

Proof:

Let $q_n = Pr\{X_n = 0), i.e., q_n$ is the probability that extinction occurs at or before the n^{th} generation. Clearly $q_n = P_n(0), q_1 = P_1(0) = P(0) = p_0$and from (2.5)

$$q_n = p(q_{n-1}) \tag{3.2}$$

If $p_0 = 0$, then $q_1 = 0, q_2 = 0, .., i.e.$,if the probability of no offspring is zero, extinction can never occur. If $\boldsymbol{p_0 = 1, then} q_1 = 1, q_2 = 1, \ldots, i.e.$, if the probability of no offspring is one then the extinction is certain to occur right the 0^{th} generation. So we confine ourselves to the case $0 < p_0 < 1$.

As $P(s)$ is a strictly increasing function of $s, q_2 = P(q_1) = P(p_0) > P(0) = q_1$. Assuming that $q_n > q_{n-1}$ we get $q_{n-1} = P(q_n) > P(q_{n-1}) = q_n$ and by induction $q_1 < q_2 < q_3$ ... The monotone increasing sequence $\{q_n\}$ is bounded above by 1. Hence q_n must have a limit $lim_{n\to\infty} = q\ (say), 0 \le q \le 1; q$ is the probability of ultimate extinction. From (3.2) it follows that q satisfies $q = P(q), i.e., q$ is a root of the equation (3.1),

$$s = P(s).$$

We now investigate further about the root. First, we show that q is the smallest positive root of (3.1). Let s_0 be an arbitrary positive root of (3.1). Then $q_1 = P(0) < P(s_0) = s_0$ and assuming that $q_m < s_0$, we get $q_{m+1} = P(q_m) < P(s_0) = s_0$ and by induction $q_n < s_0$ for all n. Thus $q = lim_{n\to\infty} \le s_0$, which implies that q is the smallest positive root of (3.2).

For this, we consider the graph of $y = P(s)\ in\ 0 \le s \le 1$; it starts with the point$(0, p_0)$ and ends with the point$(1,1)$; the curve lying entirely in the first quadrant, is convex as $P(s)$ is an increasing function. So the curve $y = P(s)$ can intersect the line $y = s$ in at most two points, one of which is the end point $(1,1), i.e.$, the equation (3.1) has at most two roots, one of which is unity. Two cases are now to be considered (see Figs. 9.1 and 9.2).

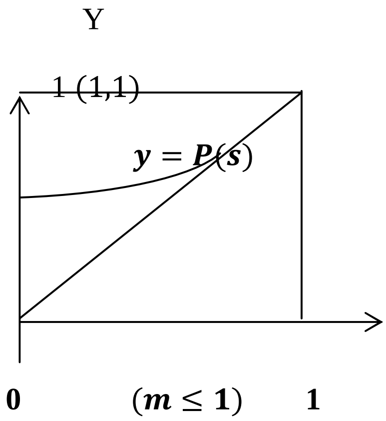
Case I.

The curve $y = P(s)$ lies entirely above the line $y = s$; in this case(1,1)is the only point of intersection, i. e., unity is the unique root of $s = P(s)$ so that Thus $q = lim_{n\to\infty} q_n = 1$. Then

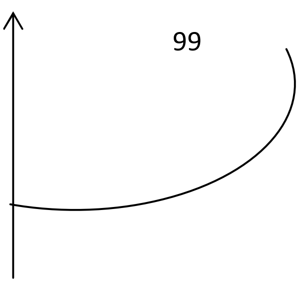
$$P(1) - P(s) = 1 - P(s) \le 1 - s.$$

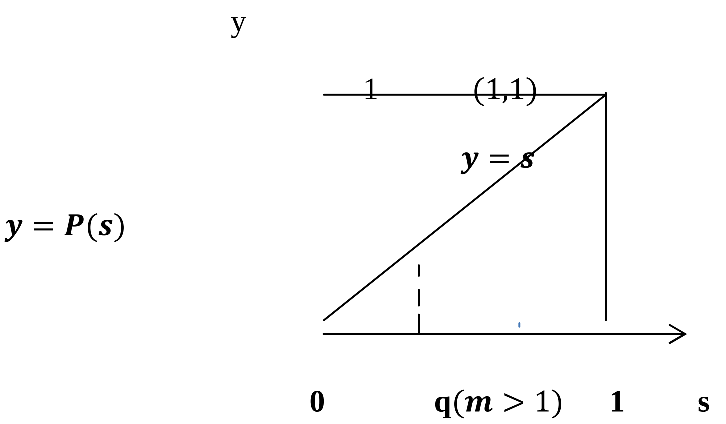
So that $$lim_{s\to 0} \frac{P(1)-P(s)}{1-s} \le 1, i.e., P'(1) \le 1.$$

Thus $$lim_{n\to\infty} q_n = 1\ when\ P'(1) = m \le 1.$$



Fig, 9.1 Graphical determination of the roots of $s = P(\boldsymbol{s})[\boldsymbol{m \le 1}]$





Fig, 9.1 Graphical determination of the roots of $s = P(\boldsymbol{s})[\boldsymbol{m} > 1]$

Case II.

The curve $y = P(s)$ intersects $y = s$ at another point$\big(\delta, P(\delta)\big)$ $such\ that\ \delta = P(\delta), \delta < 1, i.e.$, there is another root of (3.1) namely $\delta < 1$; the curve $y = P(s)$, being convex, lies below the line $y = s\ in\ (\delta, 1)$, and above $y = s\ in\ (0, \delta), i.e., P(s) < s\ in\ \delta < s < 1\ and\ P(s) > s\ in\ 0 < s < \delta$.
Then $q_1 = P(0) < P(\delta) = \delta$ and assuming that $q_m < \delta, we\ get\ q_{m+1} = P(q_m) < P(\delta) = \delta$and by induction $q_n < \delta$ for all n.

Hence $lim_{n\to\infty} q_n = \delta, so\ that\ q = \delta < 1$.

Now by the mean value theorem considered in the interval $[\delta, 1]$, there is a value $\xi\ in\ \delta < \xi < 1$ such that $P'(\xi) = \frac{P(1)-P(\delta)}{1-\delta} = 1$ and as the derivative is monotone $P'(1) > 1$.

Thus we find that q is the root less than unity of (3.1) when $m = P'(1) > 1$.

We have thus proved the theorem completely.

9.1.3 Note:

That q is a root of $s = P(s)$ can also be seen by conditioning on the number of off-springs of the first generation. We have

$$q = \Pr\ (ultimate\ extinction)$$

$$= \sum_{k=0}^{\infty} \Pr\{ultimate\ extinction \mid X_t =$$

$k. PrX1=k.$

Given that $X_1 = k$, the population will extinct iff each of the k families started by members of the 1^{st} generation becomes extinct. It is assumed that families behave independently; hence

$$\Pr\{ultimate\ extinction \mid X_1 = k\} = q^k.$$

Thus $$q = \sum_{k=0}^{\infty} q^k p_k = P(q).$$

Theorem 9.1.4.

Whatever be the value of $E(X_n) = m, we\ have, as\ n \to \infty, \lim \Pr\{X_n = 0\} = q$ and $\lim Pr\{X_n = k\} = 0$, for any finite positive integral k.

Proof: We first show that $lim_{n\to\infty} P_n(s) = q$, from which the above result will follow.

Consider the case $m \le 1, when\ P(s) = s$ has the unique root $q = 1. In\ 0 \le s \le q,\ P(s) \le P(q) = q, and$

$P_2(s) \le P_2(q) = P(P_1(P)) = P(q) = q$. Assuming that $P_m(s) \le q, we\ get\ P_{m+1}(s) \le (s) \le q$, and by induction $P_n(s) \le q$ for all n. Again $P_n(s) \ge P_n(0) = q_n; thus\ q_n \le P_n(s) \le q$.

Hence $$lim_{n\to\infty} P_n(s) = q, \qquad 0 \le s \le q.$$

Consider the case $m > 1$, when q is the root less than 1 of $P(x) = x. In\ q < s < 1$, the curve $y = P(x)$ lies below the line $y = x, and\ q < p(s) < s < 1$. Again$P_2(s) = P(P_1(s)) > P(q) = q$. Assuming that $P_m(s) > q$, we get $P_{m+1}(s) > q$, so that, by induction $P_n(s) > q$ for all n. Again $P_2(s) = P_1(P(s)) < P_1(s)$ and assuming that $P_m(s) < P_{m-1}(s)$, we get $P_n(s) < P_{n-1}(s)$for all n.

Thus in $$q < s < 1,$$

$$q < P_n(s) < P_{n-1}(s) < \cdots$$

So that $$lim_{n\to\infty} P_n(s) \ge q.$$

Suppose, if possible, that $lim_{n\to\infty} P_n(s) = \alpha > q$, then$P(\alpha) < \alpha, and\ lim_{n\to\infty} P_{n+1}(s) = lim_{n+\infty}\ P(P_n(s))$,

and we get a contradiction which is due to our supposition that $\alpha > q$. Thus

$$lim_{n\to\infty} P_n(s) = q.$$

So, whatever be the value of $E(X_1) = m, lim_{n\to\infty} P_n(s) = q$ is independent of s. In other words, for all $s < 1$,

$$lim_{n\to\infty} \sum_{k=0}^{\infty} Pr\{X_n = k\}\, s^k = q.$$

This implies that the coefficients of s^k for $k \geq 1$(except, possibly, for infinitely large k) all tend to 0, while the constant term tends to q.

Thus, as $n \to \infty$,

$Pr\{X_n = k\} \to 0$, for finite positive integral k,

and $Pr\{X_n = 0\} \to q$.

Since $P_n(1) = 1$, we have, as $n \to \infty$,

$$Pr\{X_n \to \infty\} \to 1 - q.$$

9.1.5 Note:

The above result also follows from the general theory of Markov chains applied to the chain $\{X_n\}$, for which each of the states $k = 1, 2, 3, \ldots$ is transient while the state 0 is absorbing. We shall now consider another interesting result.

Theorem 9.1.6

We have, for $r, n = 0, 1, 2, \ldots$

$$E\{X_{n+r} \mid X_n\} = X_n m^{r.} \quad (3.4)$$

Proof: For $r = 1$ and $n = 0, 1, 2, \ldots$, we have

$$E\{X_{n+r} \mid X_n\} = E\left\{\sum_{k=1}^{X_n} \varsigma_k \mid X_n\right\} = \sum_{k=1}^{X_n} E\{\varsigma_k\}$$

$$mX_n.$$

Assume that (3.4) holds for $r = k; then\ E\{X_{n+r} \mid X_n\} = X_n m^k$. Nothing the Markov nature of $\{X_n\}$, we get

$$E\{X_{n+k+1} \mid X_n\} = E[E\{X_{n+k+1} \mid X_{n+k},\ X_{n+k-1}, \ldots, X_n\} \mid X_n]$$

$$= E[E\{X_{n+k+1} \mid X_{n+k}\} \mid X_n]$$

$$= E[mX_{n+k} \mid X_n] = m(X_n m^k)$$

$$= X_n m^{k+1},$$

So that the result holds for $r = k + 1$.

Thus by induction, we have the result.

9.1.7 Asymptotic Distribution of X_n

Another variable of interest is $W_n = \frac{X_n}{m^n}$, $n = 0, 1, 2, \dots$; $\{W_n\}$ forms a Markov chain. We have $E\{W_n\} = 1$ and for $m > 1$,

$$E\{W_n^2\} = \frac{1}{m^{2n}}\ E\{X_n^2\} = \frac{1}{m^{2n}}\left\{m^{2n} + \frac{m^{n-1}(m^n-1)\sigma^2}{(m-1)}\right\}$$

$$= 1 + \frac{\sigma^2}{m^2-m}\ (1 - m^{-n}).$$

Dividing both sides of (3.4) by m^{n+r}, we get

(3.5)

$$E\{W_{n+r} \mid W_n\} = W_n$$

and since $\{W_n\}$ is also a Markov chain,

$$E\{W_{n+r} \mid W_n, W_{n-1}, \dots, W_0\} = E\{W_{n+r} \mid W_n\} = W_n.$$

(3.6)

It follows that $(W_n, n \geq 0)$ is also a martingale; further W_n being non – negative, is a non – negative martingale.

Limiting distribution of X_n

One can now apply the martingale convergence theorem for the convergence of W. Thus we get that, with probability one,

$lim_{n\to\infty} W_n$ exists and is finite.

Two cases arise:

(i) $m > 1$:

then in order that W_n converges, X_n must go to ∞ at an exponentially fast rate of n

$(so\ that\ \frac{X_n}{m^n} \to a\ finite\ limit).$

$(ii)\ \ m \leq 1$:

That as $m_n \to 0\ as\ n \to \infty$.

This implies ultimate extinction in the subcritical and critical cases.

9. 1. 8 Examples

We now consider some simple example; to fix the ideas numerical values have been taken.

Example 3(a).

Let $p_k, k = 0, 1, 2$ be the probability that an individual in a generation generates k offsprings. Then $P(s) = p_0 + p_1 s + p_2 s^2, and\ p_2(s), p_3(s)$ can be calculated by simple algebra. The probability of extinction is one if $m \le 1; if\ m > 1$, it is given by the root less than 1 of $s = P(s)$. Suppose that $p_0 = 2/3\,, p_1 = 1/6,$ and $p_2 = 1/6\,; then\ m = 1/2 < 1.$ The equation $s = P(s)$ becomes $s^2 - 5s + 4 = 0$ with roots 1 and 4; the probability of extinction is 1. Suppose that $p_0 = 1/4\,, p_1 = 1/4, p_2 = 1/2;$ then $m = 5/4 > 1$; the equation $s = P(s)$ has the roots $1/2$ and 1. The root $1/2$ gives the probability of extinction. Note that the probability of extinction is $\frac{p_0}{p_2}$ or 1 according as $p_0 < p_2$ or $p_0 \ge p_2$ and also that $p_0 < (or\ \ge) p_2\ iff\ m > (or\ \le)$

Example 3(b).

Let the probability distribution of the number of off – springs generated by an individual in a generation be Poisson with mena $\lambda\ i.e. P(s) = e^{\lambda(s-1)}$. It can be easily seen that the graph of $P(s)$ in $0 \le s \le 1\left(i.e., between\ the\ points\ \left(0, e^{-\lambda}\right) and\ (1,1)\right)$ is convex, and that the curve of $y = P(s)$ always lies above$y = s\ when\ \lambda \le 1$, there being no other root of $s = P(s)$ except unity in (0, 1); the probability of extinction is then 1. When $\lambda > 1$, the curve $y = P(s)$ intersects $y = s$ in another point whose s – coordinate has a value < 1 and the probability of extinction will be this value of s. For example, if $\lambda = 2$, it can be seen that $s = e^{2(s-1)}$ has a root approximately equal to 0.2 which is smaller than 1, and the probability of extinction is $q = 0.2$.

Example 3(c).

Let the distribution of the number of off – springs be geometric with $p_k = b(1-b)^k, k = 0, 1, 2, \ldots (0 < b < 1)$. Then $m = \frac{(1-b)}{b}\ and\ P(s) = \frac{b}{(1-s(1-b))}$. The equation $s = P(s)$ has the roots 1 and $\frac{b}{(1-b)}$. If $m \le 1$, then the probability of extinction is 1; if $m > 1$, the root $\frac{b}{(1-b)} < 1$, and the probability of extinction is equal to the root $\frac{b}{(1-b)}$.

Example 3(d).

Let $p_k = bc^{k-1}, k = 1, 2, \ldots, 0 < b, c, b + c < 1 \; and \; p_0 = 1 - \sum_{k=1}^{\infty} p_k$. Then $m = \frac{b}{(1-c)^2}$.

We have

$$P(s) = 1 - \frac{b}{1-c} + \frac{bs}{1-cs}. \quad (3.7)$$

The quadratic equation $s = P(s)$ has the roots

$$1 \; and \; \frac{1-(b+c)}{c(1-c)} = s_0(say).$$

If $m = 1, then \; s_0 = 1$ and the probability of extinction is 1; if $m > 1, s_0 < 1$, and the probability of extinction is $q = s_0 (< 1)$.

This model was applied in a series of interesting papers by Lotka to find out the probability of extinction for American male lines of descent. The values estimated by him (in 1939) from census figures of 1920 give b = 0.2126, c = 0.5893 (m = $1.25 > 1$) and the probability of extinction $q = s_0 = 0.819$.

9.1.9 Note:

It is not always possible to put the generating functions $P_n(s)$ in closed form. The generating functions $P(s)$ obtained in Examples 3(c) and 3(d) are of interesting forms: they may be considered as particular cases of the more general fractional linear form (or general bilinear form)

$$P(s) = \frac{\alpha + \beta s}{\gamma + \delta s}, \alpha\delta - \beta\gamma \neq 0.$$

When $P(s)$ is the above form, $P_n(s)$ is also of the same form

$$P_n(s) = \frac{\alpha_n + \beta_n s}{\gamma_n + \delta_n s}$$

Where $\alpha_n, \; \beta_n, \; \gamma_n, \; \delta_n$ are functions of $\alpha, \beta, \gamma, \delta$.

Further, it may be noted that the equation $s = P(s)$ (where P(s) is of fractional linear form) has two finite solutions 1 and $s_0 <, =, > or \; 1$ according as $m = P'(1) >, =, or < 1$.

Example 3(e).

$P_n(s)$ for Lotka's model considered in Example 3(d) above. For any two points u, v, we get

$$\frac{P(s)-P(u)}{P(s)-P(v)} = \frac{s-u}{s-v} \cdot \frac{1-cv}{1-cu}.$$

Put $u = s_0, v = 1, then\ P(s_0) = s_0, P(v) = 1, so\ that$

$$\frac{P(s)-s_0}{P(s)-1} = \frac{s-s_0}{s-1} \cdot \frac{1-c}{1-cs_0}$$

Whence $$\frac{1-c}{1-cs_0} = \left\{\frac{P(s)-s_0}{s-s_0}\right\} / \left\{\frac{P(s)-1}{(s-1)}\right\}.$$

Let $m \neq 1$, then taking limits of the right hand side as $s \to 1$, we get

$$\frac{1-c}{1-cs_0} = \frac{1}{m}.$$

Hence $$\frac{P(s)-s_0}{P(s)-1} = \left(\frac{1}{m}\right)\frac{s-s_0}{s-1}.$$

Thus $$\frac{P_2(s)-s_0}{P_2(s)-1} = \frac{P(P(s))-s_0}{P(P(s))-1} = \left(\frac{1}{m}\right)\frac{P(s)-s_0}{P(s)-1}$$

$$= \left(\frac{1}{m^2}\right)\frac{s-s_0}{s-1}$$

and on iteration

$$\frac{P_n(s)-s_0}{P_n(s)-1} = \left(\frac{1}{m^n}\right)\frac{s-s_0}{s-1}, n = 1, 2, \ldots$$

Solving for $P_n(s)$, we get

$$P_n(s) = 1 - m^n\left(\frac{1-s_0}{m^n-s_0}\right) + \frac{m^n\left(\frac{1-s_0}{m^n-s_0}\right)^2 s}{1-\left(\frac{m^n-1}{m^n-s_0}\right)s}, m \neq 1. \quad (3.8)$$

If $m = 1, then\ s_0 = 1\ and\ P(s) = \frac{c+(1-2c)s}{1-cs}$

and $$P(s) = \frac{nc+(1-c-ns)s}{(1-c+ns)-ncs}. \quad (3.9)$$

Limiting Results

Suppose that $m = 1, then\ \sigma^2 = \frac{2c}{(1-c)}$,

$$nPr\{X_n > 0\} = n\{1 - P_n(0)\} = \frac{n(1-c)}{1-c+nc}$$

and $lim_{n\to\infty} nPr\{X_n > 0\} = \frac{1-c}{2} = \frac{2}{\sigma^2}$ (see Theorem 9.8(a))

Suppose that $m < 1, then\ s_0 > 1$ and

$$lim_{n\to\infty} m^{-n} Pr\{X_n > 0\} = lim_{n\to\infty} m^{-n}\{1 - P_n(0)\}$$

$$= lim_{n\to\infty} \frac{1-s_0}{m^n - s_0} = \frac{s_0 - 1}{s_0}.$$

Again $\sum_k Pr\{X_n = k \mid X_n > 0\} s^k = \frac{P_n(s) - P_n(0)}{1 - P_n(0)}$

$$= 1 - \frac{1 - P_n(s)}{1 - P_n(0)}$$

$$= \frac{\left(\frac{1-s_0}{m^n - s_0}\right) s}{1 - \left(\frac{m^n - 1}{m^n - s_0}\right) s}.$$

Thus $lim_{n\to\infty} \sum_k Pr\{X_n = k \mid X_n > 0\} s^k = s\Big(1 -$

1s01−1s0−1

and $lim_{n\to\infty} Pr\{X_n = k \mid X_n > 0\} = \left(1 - \frac{1}{s_0}\right)\left(\frac{1}{s_0}\right)^{k-1},\ k =$ 1, 2, ….

In other words, for large n, the distribution of $\{X_n\}$, given $X_n > 0$, is geometric with mean $\frac{s_0}{s_0 - 1}$ and p. g. f.

$$b(s) = \frac{(1 - 1/s_0)\, s}{1 - s/s_0}.$$

It can be easily verified that $b(s)$ satisfies the equation

$$b(P(s)) = mb(s) + 1 - m$$

(see Theorem 9.9).

9.2 DISTRIBUTION OF THE TOTAL NUMBER OF PROGENY

Let X_n denote the size of n^{th} generation, $n = 0, 1, 2, \ldots, and\ X_0 = 1$.

Then the random variable

$$Y_n = \sum_{k=0}^{n} X_k = 1 + X_1 + \cdots + X_n \qquad (4.1)$$

Denote the total number of progeny, i. e., the number of descendants up-to and including the n^{th} generation and also including the ancestor.

Theorem 9.2.1.

The p. g. f. $R_n(s)$ of Y_n satisfies the recurrence relation

$$R_n(s) = sP\big(R_{n-1}(s)\big), \quad (4.2)$$

$P(s)$ being the p. g. f. of the offspring distribution.

Proof: Let $Z_n = X_1 + \cdots + X_n$ and $G_n(s)$ be its p. g. f. Then $R_n(s) = s\, G_n(s)$. We have

$$G_n(s) = \sum_{k=0}^{\infty} Pr\{Z_n = k\} s^k.$$

Now by conditioning on the size X_1 of the 1^{st} generation, we get

$Pr\{Z_n = k\} = \sum_{i=0}^{\infty} Pr\{$ total number of descendants in the succeeding $(n-1)$ generations following the first is $k - i \mid X_1 =$ i.Pr$X1=i$.

If the process starts with one ancestor then the probability of having r descendants in succeeding m generations is the coefficient of $s^r in\ G_m(s)$; and if it starts with i ancestors then the probability of having r descendants in the succeeding m generations will be the coefficient of $s^r in\ [G_m(s)]^i$. Thus

$$Pr\{Z_n = k\} = \sum_{i=0}^{\infty} \big[coefficient\ of\ s^{k-1}\ in\ \{G_{n-1}(s)\}^i\big] p_i$$

$$= \sum_{i=0}^{\infty} p_i \big[coefficient\ of\ s^k\ in\ \{sG_{n-1}(s)\}^i\big]$$

$$= coefficient\ of\ s^k\ in\ \sum_{i=0}^{\infty} p_i \{sG_{n-1}(s)\}^i$$

$$= coefficient\ of\ s^k\ in\ P\ \big(sG_{n-1}(s)\big).$$

Thus $\quad G_n(s) = \sum_{k=0}^{\infty} Pr\,\{Z_n = k\} s^k = P\big(sG_{n-1}(s)\big),$

Whence $\quad R_n(s) = sG_n(s) = sP\big(R_{n-1}(s)\big).$

Hence the theorem.

From the recurrence relation (4.2) it is theoretically possible to calculate $R_1(s), R_1(s), \ldots$. We are however interested in the asymptotic behavior of $R_n(s)$for large n.

Theorem 9.2.2

We have

$$lim_{n\to\infty} R_n(s) = H(s)$$

Where $H(s) = \sum_{k=0}^{\infty} \rho_k s^k$ is the generating function of the sequence of non – negative numbers ρ_k,

The function $H(s)$ satisfies the relation

$$H(s) = sP\big(H(s)\big), 0 < s < 1; \quad (4.3)$$

Further, $H(s)$ is the unique root of the equation

$$t = sP(t) \quad (4.4)$$

Such that $H(s) \le \xi, where\ \xi$ is the smallest positive root of $x = P(x)$ and that

$$H(1) = \sum \sum_{k=0}^{\infty} \rho_k s^k = \xi.$$

Proof: We have, for $0 < s < 1$,

$$R_2(s) = sP\big(R_1(s)\big) < sP(s) = R_1(s)$$

and assuming that $R_m < R_{m-1}(s)$, we get

$$R_{m-1}(s) = sP(R_m(s)) < sP\big(R_{m-1}(s)\big) = R_m(s)$$

and hence by induction $R_m(s) < R_{n-1}(s)\ for\ all\ n > 0$. Thus for $s < 1, \{R_m(s)\}$ is a monotone decreasing sequence bounded below. Hence $lim_{n\to\infty} R_n(s) = H(s)$ exists. From the continuity theorem of p. g. f. 's it follows that H(s), being the limit of a sequence of p. g. f. 's, is the generating function of a sequence of non – negative numbers ρ_k such that $H(1) = \sum \rho_k \le 1$.

Taking limit of (4.2), we get

$$H(s) = sP\big(H(s)\big), \qquad 0 < s < 1,$$

i.e., for some fixed $s(in\ 0 < s < 1),\ H(s)$ is a root of the equation

$$t = sP(t)$$

For fixed $s < 1, y = sP(t)$ is a convex function of t and the graph of $y = sP(t)$ intersects the line $y = t$ in at most two points. Let ξ, be the smallest positive root of $x = P(x)$; clearly $\xi \le 1$. The function $t - sP(t)$ is negative for $t = 0$ and positive for $t = \xi$, and remains positive for values of t between ξ and 1. Thus $t = s\,P(t)$ has exactly one root between $0\ and\ \xi$ and no root between ξ and 1. The unique root of $t = s\,P(t)$ equals $H(s)$ and thus $H(s) <$

ξ. Clearly $H(1)$ is a root of $x = P(x)$ and since ξ is the smallest root of this equation, $H(1) = \xi$.Thus the theorem is completely established.

9.2.3 Note:$H(1) = 1 \; or < 1$ depending on whether $1 + X_1 + X_2 + \cdots$ is finite or not (with probability one); and $H(1) = 1 \; whenever \; m \leq 1$.

9.3 CONDITIONAL LIMIT LAWS – Kolmogrov and Yaglom

9.3.1 Critical Processes

In Example 3(e) we obtained some limiting results. We shall obtain here some general results. Consider a critical (i.e., with m=1) G. W. process. The probability of extinction is 1. Thus from Theorem 9.4, we get $Pr\{X_n \to 0\} = 1$. We also have $var(X_n) = n\sigma^2 \to \infty$.The distribution of X_n, gives that $X_n > 0$, is of **considerable interest.**

9.3.2 Lemma.

For a G. W. process with $m = 1 \; and \; \sigma^2 < \infty$, we have

$$lim_{n\to\infty} \frac{1}{n}\left\{\frac{1}{1-P_n(s)} - \frac{1}{1-s}\right\} \to \frac{\sigma^2}{2}$$

(5.1)

Uniformly in $0 \leq s < 1$.

Proof:

Let $0 \leq s < 1$ and $P'''(1) < \infty$. Using Taylor's expansion of $P(s)$ in the neighbourhood of 1,

We get $$P(s) = s + \frac{\sigma^2}{2}(1-s)^2 + r(s)(1-s)^2,$$

(5.2)

Where $r(s) \to 0 as \; s \to 1$.

Thus $$\frac{1}{1-P(s)} - \frac{1}{1-s} = \frac{P(s)-s}{(1-s)\,(P(s)-s)}$$

$$= \frac{1-s}{1-P(s)}\left\{\frac{\sigma^2}{2} + r(s)\right\}$$

$$= \left\{\frac{\sigma^2}{2} + r(s)\right\}\left[1 - (1-s)\left\{\frac{\sigma^2}{2} + \right.\right.$$

r(s)−1

$$= \frac{\sigma^2}{2} + R(s),$$

(5.3)

Where $R(s) \to 0$ as $s \to 1, and$ R is bounded. Again using (4.6), we get

$$P_2(s) = P\big(P(s)\big) = P(s) + \frac{\sigma^2}{2}(1 - P(s))^2 + r\big(P(s)\big)(1 - P(s))^2$$

So that $$\frac{1}{1-P_2(s)} - \frac{1}{1-P(s)} = \frac{\sigma^2}{2} + R(P(s)) \qquad (5.4)$$

and $$\frac{1}{2}\left\{\frac{1}{1-P_2(s)} - \frac{1}{1-s}\right\} = \frac{\sigma^2}{2} + \frac{1}{2}\big\{R(s) + R\big(P(s)\big)\big\}.$$

Iterating one gets $$\frac{1}{n}\left\{\frac{1}{1-P_2(s)} - \frac{1}{1-s}\right\} = \frac{\sigma^2}{2} + \frac{1}{n}\sum_{k=0}^{n-1} R\big(P_k(s)\big).$$

Since $P_n(0) \le P_n(s) \le 1 and$ $P_n(0) \to 1$from the left, the convergence of $P_n(s) \to 1$ is uniform. Hence the lemma.

We shall now use the lemma to establish the following interesting limit laws **(of which (a) is due to Kolmogorov and (b), (c) are due to Yaglom).**

Theorem 9.3.3

If $m = 1,\ \sigma^2 < \infty$, then

$$(a) lim_{n\to\infty}\, n\, Pr\{X_n > 0\} = \frac{2}{\sigma^2}$$

$$(b) lim_{n\to\infty}\, E\left\{\frac{X_n}{n} \mid X_n > 0\right\} = \frac{\sigma^2}{2}$$

$$(c) lim_{n\to\infty}\, Pr\left\{\frac{X_n}{n} > u \mid X_n > 0\right\} = exp\left(-\frac{2u}{\sigma^2}\right), u \ge 0.$$

Proof: (a)(Kolmogrov)

We have $$nPr\{X_n > 0\} = n\{1 - P_n(0)\}$$

$$= \left[\frac{1}{n}\left\{\frac{1}{1-P_n(0)} - 1\right\} + \frac{1}{n}\right]^{-1}.$$

Thus from the lemma (taking s=0), we get

$$lim_{n\to\infty}\, n\, Pr\{X_n > 0\} = lim_{n\to\infty}\left[\frac{\sigma^2}{2} + \frac{1}{n}\right]^{-1} = \frac{2}{\sigma^2}.$$

(b)(Yaglom)

We have $1 = E\{X_n\} = E\{X_n | X_n > 0\}. Pr\{X_n > 0\} + 0 \,. Pr\{X_n = 0\}$

So that $E\{X_n | X_n > 0\} = \frac{1}{Pr\{X_n>0\}} = \frac{n\sigma^2}{2}$, $\big(from\ (a)\big)$.

Thus $lim_{n\to\infty}\, E\left\{\frac{X_n}{n} \mid X_n > 0\right\} = \frac{\sigma^2}{2}$.

$(c)(Yalgom)$ *Let* $u > 0, and$ $df(u) = Pr\left\{u \le \frac{X_n}{n} < u + du \mid Xn>0;\right.$ (5.5)

Then taking L. T. (see equation (3.1a) Chapter1), we get

$$\int_0^\infty exp\,(-\alpha u)dF(u) = E\left\{exp\left(-\frac{X_n}{n}\right) \mid X_n > 0\right\}. \quad (5.6)$$

Now $E\left\{exp\left(-\frac{X_n}{n}\right)\right\} = E\left\{exp\left(-\frac{X_n}{n}\right) \mid X_n > 0\right\}.\ Pr\{X_n > 0\} + 1\,.\ Pr\{X_n = 0\}$

And since $P_n(s) = E\{S^{X_n}\}$ is the p. g. f. of X_n, we get

$$P_n\left(exp\left(-\frac{\alpha}{n}\right)\right) = E\left\{exp\left(-\frac{\alpha X_n}{n}\right) \mid X_n > 0\right\}\{1 - P_n(0)\} + P_n(0).$$

Thus

$$E\left\{exp\left(-\frac{\alpha X_n}{n}\right) \mid X_n > 0\right\} = \frac{P_n\big(exp(-\alpha/n)\big) - P_n(0)}{1 - P_n(0)}$$

$$= 1 - \frac{1 - P_n\big(exp(-\alpha/n)\big)}{1 - P_n(0)}. \quad (5.7)$$

Now as $n \to \infty$

$$n\left\{1 - P_n\left(exp\left(-\frac{\alpha}{n}\right)\right)\right\} \to \frac{2}{\sigma^2}\ (from(a))$$

and from the basic lemma (because of uniform convergence),we get

$$\frac{1}{n\{1-P_n(exp(-\alpha/n))\}} = \frac{1}{n}\left\{\frac{1}{1-P_n(exp(-\alpha/n))} - \frac{1}{1-exp(-\alpha/n)}\right\} + \frac{1/n}{1-exp(-\alpha/n)}$$

$$\to \frac{\sigma^2}{2} + \frac{1}{\alpha}.$$

Thus from (5.6) and (5.7), we get, $n \to \infty$

$$\int_0^\infty exp\,(-\alpha u) dF(u) \to 1 - \frac{\sigma^2/2}{\sigma^2/2+1/\alpha}.$$

$$= \frac{1}{1+\alpha\sigma^2/2}.$$

Since L. T. of $\frac{2}{\sigma^2}\, exp\left(-\frac{2u}{\sigma^2}\right)$ is $\frac{1}{1+\alpha\sigma^2/2}$,

We have $lim_{n\to\infty}\, Pr\left\{u \le \frac{X_n}{n} < u + du \mid X_n > 0\right\} = \frac{2}{\sigma^2}\, exp\left(-\frac{2u}{\sigma^2}\right)$,

Which establishes the exponential limit law. ☐

9. 3. 4 Subcritical Processes

Theorem 9.3.4*(Yaglom's Theorem).*For a Galton – Watson process with $m < 1$,

$$lim_{n\to\infty}\, Pr\,\{X_n = j \mid X_n > 0\} = b_j, \qquad j = 1, 2, \ldots \qquad (5.8)$$

Exists, and$\{b_j\}$ gives a probability distribution whose p. g. f.

$B(s) = \sum_{j=1}^{\infty} b_j s^j$ satisfies the equation

$$B\big(P(s)\big) = mB(s) + 1 - m \qquad (5.9)$$

i. e., $1 - B\big(P(s)\big) = m\big(1 - B(s)\big)$.

Further $\sum_{j=-1}^{\infty} jb_j = 1/\phi(0), where\ \phi(0) = lim_{n\to\infty}\, Pr\{X_n > 0\}/m^n$.

Proof: Using Taylor's expansion around $s = 1$, we get

$$P(s) = 1 - m(1-s) + (1-s)\, r(s),\ \ 0 \le s \le 1 \qquad (5.10)$$

or, $\frac{1-P(s)}{1-s} = m - r(s)$. (5.10a)

Consider the function $r(s)\ in\ 0 \le s \le 1$; we have $r(0) = m - (1 - p_0) \ge 0$

and $lim_{s\to 1-0}\, r(s) = 0$. Further as $P(s)$ is a convex function

$$P'(s) \leq \frac{1 - P(s)}{1-s},$$

So that $$P'(s) = (1-s)^{-1}\left\{\frac{1 - P(s)}{1-s} - P'(s)\right\} \leq 0.$$

Thus $r(s)$ is monotone decreasing, and is bounded above by $m \; and \; r(s) \rightarrow 0 \; as \; s \rightarrow 1$. Replacing $s \; by \; P_{k-1}(s)$ in (5.10a), we get

$$\frac{1 - P_k(s)}{1 - P_{k-1}(s)} = m\{1 - r\,(P_{k-1}(s)/m\}.$$
(5.11)

Putting $k = 1, 2, 3, \ldots, n$ and taking products of both sides, we get

$$\frac{1 - P_n(s)}{1-s} = m^n \prod_{k=0}^{n=1}\{1 - r(P_k(s))/m\}.$$

Since $0 \leq r/m \leq 1$, the sequence $\left\{\frac{1 - P_n(s)}{m^n\,(1-s)}\right\}$ is monotone decreasing in n, and we have

$$lim_{n\rightarrow\infty} \frac{1 - P_n(s)}{m^n\,(1-s)} = \phi(s) \geq 0.$$

Putting $s = 0$, we get

$$\phi(0) = lim_{n\rightarrow\infty} \frac{1 - P_n(0)}{m^n} lim_{n\rightarrow\infty} \frac{\Pr|X_n > 0|}{m^n}.$$

Let $b_{jn} = Pr\{X_n = j |\, X_n > 0\} \; and \; B_n(s) = \sum_{j=1}^{\infty} b_{jn}\, s^j$ be the p. g. f. of $\{b_{jn}\}$. Then

$$B_n(s) = \frac{P_n(s) - P_n(0)}{1 - P_n(0)} = 1 - \frac{1 - P_n(s)}{1 - P_n(s)}$$
(5.14)

$$= 1 - (1-s)\prod_{k=0}^{n=1} \frac{1 - r(P_k(s))/m}{1 - r(P_k(0))/m}$$ (from 5.12).

Since $P_k(s) \geq P_k(0), \; and \; r(s)$ is monotone decreasing, $r((P)_k(s)) \leq r(P_k(0))$, and so each factor of the product on the r. h. s. expression is larger than 1. Thus $B_n(s)$ is monotone decreasing and tends to a limit $B(s)$ as $n \rightarrow \infty$, i. e.,

$$B(s) = \sum_{j=1}^{\infty} b_j\, s^j, \quad where \; b_j = lim_{n\rightarrow\infty} b_j n = lim_{n\rightarrow\infty} Pr\{X_n = j |\, X_n > 0\}.$$

Clearly $B(0) = 0$.

Now, $$B(P_k(0)) = lim_{n\rightarrow\infty} B_n(P_k(0))$$

$$= lim_{n\to\infty}\left\{1-\frac{1-P_n\big(P_k(0)\big)}{1-P_n(0)}\right\}$$

$$= 1 - lim_{n\to\infty}\left\{1-\frac{1-P_k\big(P_n(0)\big)}{1-P_n(0)}\right\}$$

$$= lim_{s\to 1}\frac{1-P_k(s)}{1-s} \qquad (since\ m<1,$$

$P_n(0) \to 1\ as\ n \to \infty)$

$$= 1 - m^k$$

(taking limit of (5.12) as s→ 1).

It follows that lim $B\big(P_k(0)\big) \to 1. As\ m<1, P_k(0) \to 1\ as\ k \to \infty$.

Hence $B(s) \to 1 as\ s \to 1$. Thus $B(s)$ is the p. g. f. of (b_j).

Further $\quad \sum_{j=1}^{\infty} jb_j = B'(1) = lim_{k\to\infty}\frac{1-B\big(P_k(0)\big)}{1-P_k(0)}$

$$= lim_{k\to\infty}\frac{m^k}{1-P_k(0)}$$

$$= lim_{k\to\infty}\frac{m^k}{Pr\{X_k>0\}} \to \frac{1}{\phi(0)}.$$

Again, from (5.14)

$$B_n\big(P(s)\big) = 1 - \frac{1-P_n\,(P\,(s))}{1-P_n(0)}$$

$$= 1 - \frac{1-P_{n+1}\,P\,(s)}{1-P_{n+1}(0)}\cdot\frac{1-P_{n+1}\,P\,(0)}{1-P_n(0)} \qquad (5.15)$$

$$lim_{n\to\infty}\frac{1-P_{n+1}\,P\,(s)}{1-P_{n+1}(0)} = lim_{n\to\infty}\big(1-B_{n+1}(s)\big) = 1-B(s),$$

$$lim_{n\to\infty}\frac{1-P_{n+1}\,P\,(0)}{1-P_n(0)} = lim_{s\to 1}\frac{1-P(s)}{1-s} = m.$$

Hence taking limits of both sides of (5.15), We get

$$B\big(P(s)\big) = 1 - \big(1-B_{n+1}(s)\big)\,m$$

$$= mB(s) + 1 - m.$$

Thus the theorem is proved.

9.3.5 Remarks:

1. The above simplified proof which does not involve moment restriction is due to Joffe(1967).

The equation (5.9a) is known as a modified *Schroder functional equation.* It is not always easy to obtain $B(s)$ from it for given $P(s)$. In example 3(e) we obtained $B(s)$ directly.

The mean $\sum_{k=1}^{\infty} kb_k = \frac{1}{\phi(0)}$ of the limiting distribution is finite iff

$$E\{X_1 \, log X_1\} = \sum_{k=1}^{\infty} P_k\{k \log k\} < \infty \; or \; p_0 = 1.$$

2. The limiting behavior which has customarily been studied through probability generating functions and their functional iterates, has now been studied also through the martingale convergence theorem; the latter is more revealing on the nature of the process. Sec, for example, Heyde (1970) and Grey (1980).

Before closing the discussion on G. W. processes we mention a few interesting innovations introduced in the process.

UNIT X : THE CLASSICAL GALTON AND WATSON PROCESS

10.1 THECLASSICAL GALTON – WATSON PROCESS

10.1.1 Branching Processes with Immigration

Theorem 9.4 states that for a G. W. Process, $\Pr\{X_n \to 0\} = q$ *and* $Pr\{X_n \to k=0,$ for finite k *and so* $PrXn\to\infty=1-q,q$ being the probability of extinction. Further, $q = 1$ for critical and subcritical processes. Thus left to themselves G. W. populations either die out or grow without limits, Immigration from outside into a critical or subcritical process could have stabilizing effect on the population size. Apart from this aspect, immigration by itself is interesting from the point of view of theory and applications. Galton – Watson processes with immigration often arise in applications in such areas as traffic theory, statistical mechanics, genetics, neurophysiology etc.

Consider a G. W. process with off-spring distribution $\{p_k\}(having\ p.g.f.P(s)\ and\ mean\ P'(1) = m\ )$.(The process will be called underlying G. W. process.) Suppose that at time n, I. e., at the time of birth of n^{th} generation there is an immigration of Y_n objects into the population, and that $Y_n, n = 0, 1, 2, ...$ are i. i. d random variables with p. g. f.

$$h(s) = \sum_{j=0}^{\infty} Pr\{Y_n = j\}s^j = \sum h_j s^j,$$

i. e., with probability h_j, j immigrants enter the n^{th} generation and contribute to the next generation in the same way as others already present do. The numbers of immigrants into successive generations are independent and all objects reproduce independently of each other and of the immigration process. The distribution $\{h_j\}$ will be called immigrant distribution. Let $a = h'(1)$ be the mean of this distribution.

Such a process (G. W. I.) can be denoted as

$$\{X_{n+1},\ n = 0, 1, 2, ...\}$$

Where

$$X_{n+1} = \sum_{r=1}^{X_n} \zeta_r + Y_{n+1} \qquad (5.16)$$

And $E(\zeta_r) = m, E\{Y_n\} = a.$

Let $X_{(n)}$ be the number of objects at the n^{th} generation and let

$$P_{(n)}(s) = \sum_{j=0}^{\infty} Pr\{X_{(n)} = j\} s^j$$

Be its p. g. f. The sequence ($X_{(n),} n = 0, 1, 2, ...\}$ defines a G. W. I process. The sequence is a Markov chain whose one – step transition probabilities are given by

$$p_{ij} = coeff. of\ s^j\ in\ h(s)[P(s)]^i,\ i, j \in N.$$

Clearly, $P_{(n)}(s) = h(s)P_{(n-1)}(P(s)).$

If $lim_{n\to\infty} P_{(n)}(s) = F(s)$ exists, then one gets

$$F(s) = h(s)\ F(P(s))$$

i.e., the limit, when it exists, satisfies the above functional equation.

Now the question arises : when does the limit exist, and does it define the p. g. f. of a proper probability distribution or when does $\{X_{(n),}\}$ have a proper limit distribution?

10.2. BELL-MAN HARRIS PROCESS

In the preceding section we assumed that the lifetimes of objects (particles, individuals, organisms) are exponential random variables. Here we shall generalize his further; we shall consider that the lifetimes have general, and not necessarily exponential, distributions.

Suppose that an object (ancestor) at time t = 0 initiates the process. At the end of its lifetime it produces a random number of direct descendants having offspring distribution $\{p_k\}$ (with p.g.f. P (s)).

We assume, as before, that these descendants act independently of each other and that at the end of its lifetime, each descendant produces its own descendants with the same offspring distribution $\{p_k\}$, and that the process continuous as long as objects are present. Suppose that the lifetimes of objects are i.i.d.random variables with d.f. G(which is also independent of the offspring distribution).Let { X (t) , t ≥ 0} be the number of objects alive at time t. The stochastic process { X (t) , t ≥ 0} is known as an age-dependent (or general time) branching process. Such a process is also known as a Bellman-Harris process, after Bellman and Harris who considered such a process in 1948. We shall consider here Bellman-Harris type age-dependent process. An age-dependent process is , in general ,not Markovian. For a detailed account , refer to Sankaranarayanan (1989, Chapter 4).

10.2.1 Generating Function

Theorem 10.2.1.

The generating function

$$F\,(t,s) = \sum_{k=0}^{\infty} \Pr\{\,X\,(t) = k\,\}\,s^k \qquad (8.1)$$

of an age-dependent branching process { X (t) = [1 – G(t)]s + $\int_0^t P\big(\,F(t - u, sdGu.$ (8.2)

proof: To find { X (t) = k}, we shall condition on the lifetime T at which the ancestor dies bearing I offsprings.

We have

$\Pr\{\,X\,(t) = k\,\} = \int_0^{\infty} \Pr\{\,X\,(t) = k|T = u\}\,dG(u)$

$= \int_0^{t} \Pr\{\,X\,(t) = k|T = u\}\,dG(u)$

$= \int_0^{\infty} \Pr\{\,X\,(t) = k|T = u\}\,dG(u).$

In case of the second term,u $> t$. Given that T = u, the number of objects at time t is then still I (the ancestor) and the expression under the square brackets equal δ_{1k}\{ 1 – G(t)\}.

In case of the first term, u $\le t$, the ancestor dies at time u $\le$ t, leaving I ($\epsilon\ N$) direct descendants: the probability of this is pi d G (u); and further these I descendants (who independently initiate processes at time u) leave k objects in the remaining time $t - u$: the probability of this event is equal to the coefficient of s^k in the expansion of $[F\ (t - u,\ s)]^i$ as a power series in s. Thus we have,

$\Pr\{\,X\,(t) = k\,\} = \{1 - G(t)\}\delta_{1k}$

$+ \int_0^t \sum_{i=0}^{\infty} pi\ d\ G(u)\ \{coeff.of\ s^{k\ in\ the\ expnsion\ of\ [F\ (t-u,s)]^i}\ \}$

Multipliying both sides by s^k , k = 0.1.2 . . . and summing over k, we find that the l.h.s. equals $F\ (t, s)$;

And that the first term on the r.h.s. equals $\{\,1 - G(t)\}$s and the second term equals

$$\int_0^t [\sum_{i=0}^{\infty} pi\ \{\sum_{k=0}^{\infty} [\ coeff\ of\ s^{k\ in\ (\ F\ (t-u,s)^i})]\ s^k\ \}]dG(u)$$

$= \int_0^t [\sum_{k=0}^{\infty} pi\ (\,F(t - u, s))^i\ ]dG(u) =$
$\int_0^t P\big(\,F(t - u, s)\big)dG(u).$

Hence the theorem.

UNIT – XI STOCHASTIC PROCESSES IN QUEUEING SYSTEMS

11.1 STOCHASTIC MODELS - QUEUING SYSTEMS

The queueing theory had its origin in 1909, when A. K. Erlang (1878 - 1929) published his fundamental paper relating to the study of congestion in telephone traffic. The literature on the theory of queues and on the diverse areas of its applications have grown tremendously over the years.

A queue or waiting line is formed when units (or customers, clients) needing some kind of service arrive at a service channel (or counter) that offers such facility. A queueing system can be described by the flow of units for service, forming or joining the queue, if service is not immediately available, and leaving the system after being served . The basic features which characterize a system are; (i) the input, (ii) the service mechanism, (iii) the queue discipline and (iv) the number of service channels.

By units, we mean those demanding service, e. g. customers at a bank counter or at a reservation counter, calls arriving at a telephone exchange, vehicular traffic at a traffic intersection, machines for repair before a repairman, airplanes waiting for take – off at a busy airport, merchandise waiting for shipment at a yard, computer programmes waiting to be run on a time – sharing basis etc.

The *input* describes the manner in which units arrive and join the system. The interval between two consecutive arrivals is called the inter-arrival time or interval. The system is called a delay or loss system depending on whether a unit who, on arrival, finds the service facility occupied, joins or leaves the system. The system may have either a limited or an unlimited capacity for holding units. The source from which the units come may be finite. A unit may arrive either singly or in a group.

The service mechanism describes the manner in which service is rendered. A unit may be served either singly or in a batch. The time required a unit is called the *service time.*

The queue discipline indicates the way in which the units form a queue and are served. The usual discipline is *first come first served* (FCFS) *or first in first out* (FIFO), though sometimes, other rules, such as, last come first served or random ordering before service are adopted. A more realistic service discipline, called *processor – sharing,* is considered in computer science literature; this envisages that if there are $m$jobs, each receives service at the rate of $\frac{1}{m}$.

The system may have a *single channel or a number of parallel channels* for service.

The inter-arrival and service times may be deterministic or chance – dependent. The case when both the inter-arrival and service times are deterministic is trivial. We shall be generally concerned with chance – dependent inter-arrival and service times, and the theory will be essentially stochastic. When chance – dependent, the inter-arrival times between two consecutive arrivals are assumed to be i. i. d. random variables; the service times of units are also assumed to be i. i. d. random variables. Further the two sets of

random variablesare also taken to be independent.

The mean arrival rate, usually denoted by λ, is the mean number of arrivals per unit time. Its reciprocal is the mean of the inter-arrival time distribution. The mean service rate, usually denoted byμ, is the mean number of units served per unit time, its reciprocal being the mean service time. In a single channel system, the ratio

$a = \frac{\lambda}{\mu} = \frac{arrival\ rate}{service\ rate} = \frac{mean\ service\ time}{mean\ interarrival\ time}$Is called the *offered load or traffic intensity*. Though dimensionless, it is expressed in *erlangs.* It can be seenthat if $\lambda > \mu$, then the queue size will go to infinity. The quantity $\rho = a/c$ is called *carried load.*

11. 2 QUEUINGPROCESSES

The following random variables or families of random variables that arise in the study provide important measures of performance and effectiveness of a stochastic queuing system.

1. The number $N(t)$ in the system at time t, i. e. the number at time t waiting in the queue including those being served, if any.

2. The busy period which means the duration of the interval from the moment the service commences with arrival of an unit at an empty counter to the moment the server becomes free for the first time.

3. The waiting time in the queue, i. e. the duration of time a unit has to spend in the queue; also the waiting time W_n of the n^{th} arrival.

4. The virtual waiting time $W(t), i.e.$ the interval of time a unit would have to wait in the queue, were it to arrive at the instant t.

One needs to have their complete probabilistic description. It is clear that $\{N(t),\ t \geq 0\}, \{W(t),\ t \geq 0\}, \{W_n,\ n \geq 0\}$ are stochastic processes, the first two being in continuous time and the third one in discrete time. It will be seen that some of the queuing processes that we would come across are Markovian and some are semi – Markovian. From some of the non – Markovian processes that arise, Markov chains can be extracted at suitable regeneration points and semi – Markovian processes can be constructed there from. The theory of Markov chain and semi – Markov processes thus plays an important role in the study of queuing processes.

11.2.1 Notation

A very convenient notation designed by Kendall to denote queuing system has been universally accepted and used. It consists of a three – part descriptor A/B/C, where the first and second symbols denote the inter-arrival and service time distributions respectively, and third denotes the number of channels or servers. A and B usually take one of the following symbols:

M : for exponential (Markovian) distribution

E_k : for Erlang – k distribution

G : for arbitrary distribution

D : for fixed (Deterministic) interval

Thus, by an $M/G/1$ system is meant a single channel queuing system having exponential inter-arrival time distribution and arbitrary service time distribution. By M/G/1/k is meant the same system with the fourth descriptor R denoting that the system has a limited holding capacity k.

11.2.2 Steady State Distribution

$N(t)$, the number in the system at time t and its probability distribution, denoted by

$$p_n(t) = Pr\{N(t) = n \mid N(0) = .\}$$

are both time dependent. For a complete description of the queuing process we need consider transient or time – dependent solutions. It is often difficult to obtain such solutions. Further, in many practical situations, one needs to know the behaviour in steady state, i. e. when the system reaches an equilibrium state, after being in operation for a pretty long time.

It is easier and convenient to determine

$$p_n = \lim p_n(t) \qquad as\ t \to \infty$$

Provided the limit exists. It is necessary to know the condition for the existence of the limit in the first place. This will be discussed in due course. When the limit exists, it is said that the system has reached *equilibrium or steady state* and the problem then boils down to finding the steady state solutions.

11.2.3 Little's Formula

There are certain useful statements and relationships in queuing theory which holds under fairly general conditions. Though rigorous mathematical proofs of such relations are somewhat complicated, intuitive and heuristic proofs are simple enough and have been known for long. It has been argued also that conservation methods could very well be applied to supply proofs of some of these relations. Conservation principles have played a fundamental role in physical and engineering science as well as in economics etc. Similar principles may perhaps be applied in obtaining relations for queuing system in steady state. Some such relations are given below. The most important one is

$$L = \lambda\, W$$

Where λ is the arrival rate, L is the expected number of units in the system and W is the expected waiting time in the system in steady state. A rigorous of the relation has been given by Little(1961) and so the relation is known as Little's formula. This result, of great generality, is independent of the form of inter-arrival and service time distributions, and holds under some very general conditions.

Denote the expected number in the queue and the expected waiting time in the queue in steady state by $L_Q\ and\ W_Q$ respectively . These are related by a similar formula:

$$L_Q = \lambda\, W_Q.$$

11.3 TRANSIENTBEHAVIOUR

In this section we consider the transient behavior of three specific queueing systems,namely,

$M\ /\ M\ /\ 1\ /\ 1$ (no one allowed to wait), $M\ /\ M\ /\ 1\ /\ \infty$. This discussion I restricted to these two models, Since the mathematics becomes extremely complicated with the slightest relaxation of Poisson – exponential assumptions, and it is our feeling that the exhibition of some fairly simple results is sufficient for our purposes. Even these three transient derivations vary greatly in difficulty. The $M\ /M\ /\ 1\ /\ 1$ solution can be found fairly easily, but the problem becomes much more complicated when the restriction on waiting room is relaxed, or multiple servers are considered

11.3.1 Transient Behavior of $M\ /\ M\ /\ 1\ /\ 1$

The derivation of the transient probabilities $\{p_{n\,(t)\}}\}$ that an arbitrary time t there are n customers in a single-channel system with Poisson input, exponential service, and no waiting room is a straightforward procedure, since $p_n\ \ (t) = 0\ for$ all n > 1. It begins in the usual fashion from the birth-death differential equations as given by , with Å_0 = Å, , Å_n = 0, n > 0, and $\mu_1\ = \mu$:

$$\frac{dp_1(t)}{dt}\ =\ -\mu p_1(t) +\ \text{Å}p_0(t),$$

$$\frac{dp_0(t)}{dt} =\ -\text{Å}p_0(t) +\ \mu p_1(t) \qquad (2.70)$$

These differential-difference equations can be solved easily in view of the fact that it is always true that

$$p_0(t) +\ p_1(t) = 1.$$

Hence (2.70) is equivalent to

$$\frac{dp_1(t)}{dt}\ \equiv\ p_1'(t) =\ -\mu p_1(t) +\ \text{Å}[1 -\ p_1(t)].$$

So

$$p_1{}'(t) + (\text{Å} +\ \mu)p_1(t) =\ \text{Å}.$$

This is an ordinary first-order linear differential equation with constant coefficients. Its solution can be obtained from the discussion in Section 1.7 as

$$p_1(t)\ = Ce^{-(\text{Å}+\mu)t}\ +\ \frac{\text{Å}}{\text{Å}+\mu}$$

To determine C , we use the bountry value of $p_1(t)$ at t=0,which is $p_1(o)$. Thus

$$\text{C} = p_1(0) -\ \frac{\text{Å}}{\text{Å}+\mu'}$$

and consequently

$$p_1(t) = \frac{Å}{Å + \mu}\left(1 - e^{-(Å+\mu)^t}\right) + p_1(0)e^{-(Å+\mu)t},$$

$$p_0(t) = \frac{\mu}{Å + \mu}\left(1 - e^{-(Å+\mu)^t}\right) + p_0(0)e^{-(Å+\mu)t},$$

Since $p_0(t) = 1 - p_1(t)$ for all t.

The stationary solution can be found directly from (2.70)in the usual way by letting the derivatives equal zero and then, using the fact that $p_0 + p_1 = 1$,solving for p_0 *and* $p_1(M / M / 1 / k \text{ with } K = 1)$. Also, the limiting (steady, state, equilibrium)solution can be found as the limit of the transient solution of (2.71) as t goes to ∞, we find that

$p_1 = \frac{\rho}{\rho+1}$ and $p_0 = \frac{1}{\rho+1}$

Existence of the limiting distribution is always assured, independent of the value of $\rho = Å/\mu$, and thus it is identical to the stationary distribution (to see this , put K = 1 in the p_n expression for the $M / M / 1 / K$ of section 2.5).

To get a better feel for the behavior of this queueing system for small values of time, let us graph $p_1(t)$ from (2.71). First rewrite (2.71) in the form

$$p_1(t) = p_1 + be^{-ct},$$

Where,

$$p_1(t) = \frac{Å}{Å+\mu} = \frac{\rho}{\rho+1}, \qquad b = p_1(0) - p_1, \quad and\ c = Å + \mu$$

Figure 2.9 shows a sample graph of $p_1(t)$ for a case where $b > 0(Å = 0.2, \mu = 0.4, p_1(0) = 0.7)$.

We see that $p_1(t)$ is asymptotic to p_1. In addition , if the initial probability $p_1(0)$ equals the stationary probability p_1, then b = 0 and $p_1(t)$ equals the constant $p_1$for all t. In other words , the queueing process can be translated into steady state at any time by starting the process in equilibrium. This property is , in fact, true for any ergodic queueing system, independent of any assumptions about its parameters.

11.3.2 Transient Behavior of $M / M / 1 / \infty$

The transient derivation for $M / M / 1 / \infty$ is quite a complicated procedure, so presentation of it is in outline form only. A more complete picture of the details can be found in Gross and Harris (1985) and Saaty (1961). The solution of this problem postdated that of the basic Erlang work by nearly half a century, with the first published solution due to Ledermann and Reuter (1954), in which they used spectral analysis for the general birth-death process. In the same year, an additional paper appeared on the solution of this problem by Bailey (1954), and later one by Champernowne (1956) Bailey's approach to the time-dependent problem was via generating functions for

the partial differential equation, and champernowe's was via complex combinatorial methods. It is Bailey's approach that has been the most popular over the years, and this is basically the one we take. Remember that the key thing that makes this problem more difficult than may seem at first is that we are dealing with an infinite system of liner differential equations.

To begin, let it be assumed that the initial system size at time 0 is i. That is, if N(t) denotes the number in the system at time t, then N(0) = i. The differential-difference equations governing the system size are given in (1.30) as

$$p_n'(t) = -(\lambda + \mu)p_n(t) + \lambda\, p_{n-1}(t) + \mu\, p_{n+1}(t)\ (n > 0)$$
$$p_0'(t) = -\lambda\, p_0(t) + \mu\, p_1(t)$$

It turns out that we solve these time-dependent equations using a combination of probability generating functions , partial differential equations, and Laplace transforms.

Define

$$P(z,t) = \sum_{n=0}^{\infty} p_n(t) z^n \ \text{(z –complex)}$$

Such that the summation is convergent in and on the unit circle(i.e., for |z| $\leq$) , with its Laplace transform defined as

$$\ddot{P}(z,s) = \int_0^{\infty} e^{-st} P(z,t)dt \ (\ Re(s) > 0)$$

After the generating function is formed from(2.72) – it is found when the Laplace transform is taken that

$\ddot{P}(Z,s) = \frac{z^{i+1} - \mu(1-z)\overline{\overline{p_0}}(s)}{(\lambda+\mu+s)z - \mu - \lambda z^2}$, where $\overline{\overline{p_0}}$ (s) is the Laplace transform of p_0(t).

Since the Laplace transform $\bar{p}(z,s)$ converges in the region |z| ≤ 1, Re s > o,

Wherever the denominator of the right-hand side of (2.73) has zeros in that region, so must the numerator. This fact is henceforth ued to evaluate $\overline{\overline{p_0}}$(s). The denominator has two zeros, since it is quadratic in z and they are(as functions of s)

$$z_1 = \frac{\lambda + \mu + s - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}$$
$$z_2 = \frac{\lambda + \mu + s + \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}$$

Where, the square root is taken so that its real part is positive. It is clear that $|z_1| < |z_2|, z_1 + z_2 = (\lambda + \mu + s)/\lambda$ and $z_1 z_2 = \mu/\lambda$. The completion of derivation is by the use of Rouches theorem in complex analysis.

BLOCK IV

UNIT XII BIRTH AND DEATH PROCESSES IN QUEUING THEORY

12. 1 Birth and Death Processes

Let us first consider a birth and death process with state dependent birth and death rates λ_n and μ_n respectively. Let$N(t)$ be the number present at the instant t, and

$$p_n(t) = \Pr\{N(t) = n \mid N(0) = .\}$$

It was shown in Ch. 3 that $\{N(t),\ t \geq 0\}$ is a Markov process with denumerable state space $\{0, 1, 2, \ldots\}$ and that the forward Kolmogorov equations of the process are:

$$p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t)$$

$$p_n'(t) = -(\lambda_n + \mu_n) p_n(t) + \lambda_{n-1} p_{n-1}(t) + \mu_{n+1} p_{n+1}(t), n = 1, 2, \ldots$$

(see equations (4.4) and (4.6), Chapter 3).

We proceed to investigate the steady state solutions. Assume that such solutions exists, then $lim_{t\to\infty} p_n(t) = p_n = \Pr\{N = n\}$, Nbeing the random variable giving the number of units. Putting $p_n(t) = p_n$ and $p_n'(f) = 0$, we get from the above the following difference equations in steady state:

$$0 = -\lambda_0 p_0 + \mu_1 p_1 \qquad (2.1)$$

$$0 = -(\lambda_n + \mu_n) p_n + \lambda_{n-1} p_{n-1}(t) + \mu_{n+1} p_{n+1},\ n = 1, 2, \ldots \qquad (2.2)$$

Alternatively, we can obtain the steady state balance equations by using the rate – equality principle.

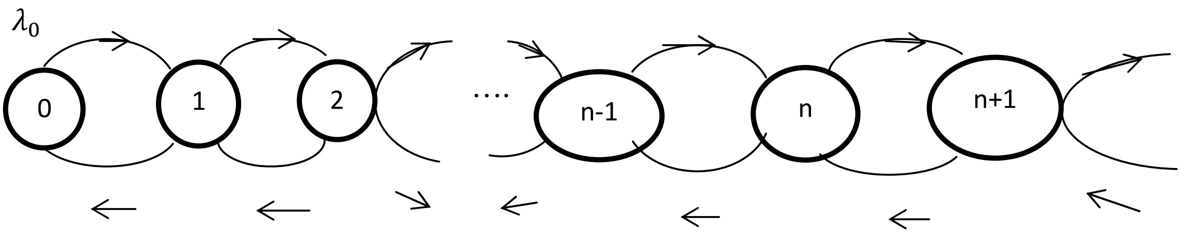


Fig. 10. 1 State – transition – rate diagram of birth – death process

Form the state – transition – rate diagram (Fig . 10. 1) it is clear that,

for $n = 0,$ $\lambda_0 p_0 = \mu_1 p_1$

And for $n > 0,$ $(\lambda_n + \mu_n)p_n + \lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1},$

(which are the equations (2.1) and (2.2)).

From (2.2), we have for $n = 1, 2..,$

$$\mu_{n+1}p_{n+1} - \lambda_n p_n = \mu_n p_n - \lambda_{n-1}p_{n-1}$$

$$= \mu_{n-1}p_{n-1} - \lambda_{n-2}p_{n-2} \qquad (putting\ n - 1\ for\ n)$$

$$= \mu_1 p_1 - \lambda_0 p_0$$

$$= 0 \quad \left(from\ (2.1)\right),$$

So that $$p_{n+1} = \frac{\lambda_n}{\mu_{n+1}} p_n$$

$$= \left(\frac{\lambda_n}{\mu_{n+1}}\right)\left(\frac{\lambda_{n-1}}{\mu_n}\right) p_{n-1}$$

… …. ….

$$= \frac{\lambda_n \lambda_{n-1} \dots \lambda_0}{\mu_{n+1}\mu_n \dots \mu_1} p_0$$

Or $$p_n = \prod_{k=1}^{n} \frac{\lambda_{k-1}}{\mu_k} p_0, \quad n = 1, 2, \dots \qquad (2.3)$$

Since $\sum_{n=1}^{\infty} p_n$ must be unity,

$$\left\{1 + \sum_{n=1}^{\infty} \prod_{k=1}^{n} \frac{\lambda_{k-1}}{\mu_k}\right\} p_0 = 1.$$
(2.4)

Hence a necessary and sufficient condition for the existence of a steady state is the convergence of the infinite series $\sum_{n=1}^{\infty} \prod_{k=1}^{n} \frac{\lambda_{k-1}}{\mu_k}$. When it converges, p_0 can be found from (2.4).

The queuing processes arising in some of the standard models can be considered as birth and death processes and the steady state solutions can be obtained easily by using the above. In the M/M/1 model considered in Sec. $10.2,\ \lambda n = \lambda, n = 0, 1, 1, \ldots, and\ \mu_n = \mu, n = 1, 2, \ldots$. Putting these values in (2.3) and (2.4) we at once get the steady state solutions.

12. 2. M/M/1 Model

In order to study the M/M/1 model , first look into the general model namely M/M/s model

Here we consider a queuing model with $s(1 \leq s \leq \infty)$ servers or channels *in parallel*and having identical input and service time distributions (as in the model M/M/1). In other words, the present model (M/M/s) considers a Poisson process with parameter λ as its input process and has, for each of the s channels, i. i. d. exponential service time distribution with mean rate μ. If $n(\leq s)$ channels are busy, the number of service completed in the whole system is given by a Poisson process with mean $n\mu$ and the time between two successive service completions is exponential with mean $1/n\mu$; whereas if $n(\leq s)$ channels are busy, the time between two successive service completions is exponential with mean $1/s\mu$. If $N(t)$ is the number present in the system at the instant t, the transition densities are as follows:

$$a_{n,n+1}\, dt \equiv Pr\{N(t+dt) = n+1 \mid N(t) = n\} = \lambda dt + o(dt)$$

$$a_{n,n+1}\, dt \equiv Pr\{N(t+dt) = n-1 \mid N(t) = n\} = \mu_n dt + o(dt)$$

$$a_{n,m}\, dt \equiv Pr\{N(t+dt) = m \mid N(t) = n\} = +o(dt),\quad m \neq n-1, n+1$$

Where $\mu_n = n\mu,\ if\ 0 \leq n \leq s$

$= s\mu,\ if\quad n \geq s.$

Thus $N(t)$ is a birth and death process with constant arrival (birth) rate $\lambda_n = \lambda$ and state – dependent service (death) rates as given in (2.5).

Let $p_n(t) = \Pr\{N(t) = n \mid N(0) = .\}$, and let the steady state solutions exist. The state – transition rate diagram is given in Fig. 10. 2.

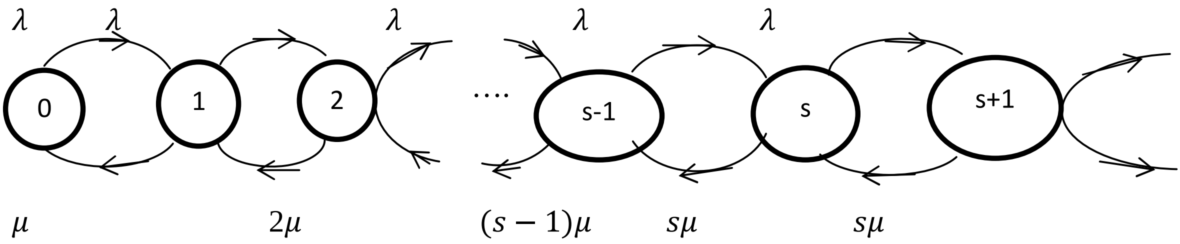


Fig. 10. 1 State transition – rate diagram of an M/M/s model

Putting the values of λ_n *and* μ_n in (2.3) and (2.4) we get p_0 *and* $p_n, n = 1, 2, \ldots,$ as follows. Denote $\lambda/s\mu$by ρ.

For $n \leq s$,

$$p_n = \frac{\lambda . \lambda \ldots \lambda}{(\mu)(2\mu)\ldots(n\mu)} p_0 = \frac{(\lambda/\mu)^n}{n!} p_0 = \frac{\lambda}{n\mu} p_{n-1}, \quad (2.6a)$$

And, for $n \geq s$

$$p_n = \frac{\lambda . \lambda \ldots n \ factors}{\{(\mu)(2\mu)\ldots(s\mu)\}\{(s\mu)\ .\ (s\mu)\ldots(n-s) factors\}} p_0$$

$$= \frac{\lambda^n}{s\,!\,\mu^s s^{n-s} \mu^{n-s}} p_0 = \frac{(\lambda/\mu)^n}{s\,!\,s^{n-s}} p_0 = p^{n-s} \frac{(\lambda/\mu)^s}{s!} p_0$$

$$= p^{n-s} p_s\,. \quad (2.6b)$$

The condition $\sum_{n=0}^{\infty} p_n = 1$ gives

$$p_0^{-1} = 1 + \sum_{n=1}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \sum_{n=s}^{\infty} \frac{(\lambda/\mu)^n}{s\,!\,s^{n-s}} = \sum_{n=0}^{s-1} \frac{(\lambda/\mu^n)}{n!} + \frac{s^s}{s\,!} \sum_{n=s}^{\infty} \left(\frac{\lambda}{s\mu}\right)^n.$$

For existence of steady state solutions, the series $\sum_{n=s}^{\infty} \left(\frac{\lambda}{s\mu}\right)^n$ must converge, and for this happen the utilization factor $\rho = \lambda/s\mu$ must be than 1. Then

$$p_0^{-1} = 1 + \sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s\,!(1-\lambda/s\mu)}. \quad (2.7)$$

Thus the distribution of N, when $\rho = \lambda/s\mu < i$ is given by (2.6), where p_0 is given by (2.7).

Note :

1. The solutions p_n satisfy the following recurrence relations

$$p_n = \frac{1}{n}\left(\frac{\lambda}{\mu}\right)^n p_{n-1}, \qquad n = 1, 2, \dots, s-1$$

$$= \frac{1}{s}\left(\frac{\lambda}{\mu}\right)^n p_{n-1} = \rho p_{n-1}, \qquad n = s, s+1, \dots.$$

For $n \le s, \{p_n\}$ behave as a Poisson distribution and for $n > s$ as a geometric distribution [when n is finite.]

2. The probability that as arriving unit has to wait is given by

$$C(s,\ \lambda/\mu) = \Pr\{N \ge s\} = \sum_{n=s}^{\infty} p_n$$

$$= \frac{(\lambda/\mu)^s}{s\,!(1-\rho)} p_0 = \frac{p_s}{1-\rho}.$$

This is known as Erlang's second (or C) formula. Extensive tables are available.

Particular Cases

12.2.1 $M/M/1\ Model$

By putting $s = 1$, we get from previuos model equation (2.7) and (2.8b)

$$p_0 = 1 - \rho = 1 - \lambda/\mu$$

$$p_1 = \rho p_0, \quad p_2 = \rho^2 p_0, \dots,$$

$$p_n = \rho^n(1-\rho), n = 0, 1, 1, \dots$$

The distribution of the number in the system N is geometric.

12.3M/M/∞ Model

Letting $s \to \infty$, we get from (2.7) and (2.8a),

$$p_0 = e^{-\lambda/\mu}, \quad p_1 = (\lambda/\mu)p_0, \qquad p_2 = \frac{1}{2}(\lambda/\mu)^2 p_0, ..$$

$$p_n = \frac{(\lambda/\mu)^n}{n!} e^{-\lambda/\mu}, \quad n = 0, 1, 2, \dots$$

The distribution of N is Poisson with parameter λ/μ.

12. 4 Model M/M/s/s: Loss Model (Due to Erlang)

This model envisages that a unit, who finds, on arrival, that all the s – channels are busy, leaves the system without waiting for service. This is called a (s - channel) *loss* system and was first investigated by Erlang. This model was also examined earlier (see Example 5(d) and 5(g), Ch.3).

For this birth and death process, we have

$$\lambda_n = \lambda, \qquad \mu_n = n\mu, \; n = 0, 1, 2, \dots s-1$$
(2.9)

$$\lambda_n = 0, \qquad \mu_n = s\mu, \; n \geq s.$$
(2.10)

The state transition – rate diagram is given in Fig. 10.3

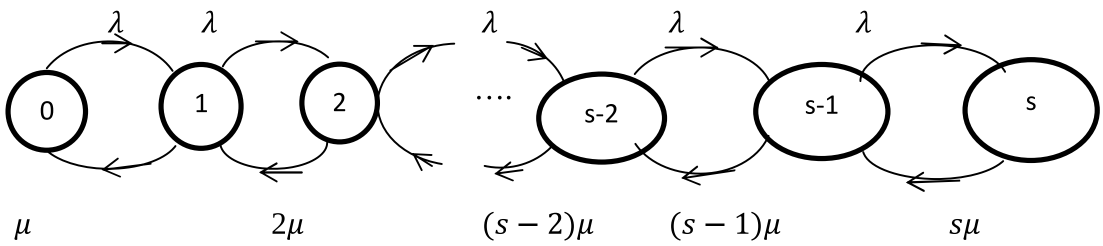


Fig. 10. 1 State transition – rate diagram for M/M/s/s model

From (2.3) and (2.4) we get the steady state probabilities

$$p_n = \frac{(\lambda/\mu)^n}{n!} p_0, \qquad n = 0, 1, \dots, s$$

And $$\left\{1 + \sum_{k=1}^{s} \frac{(\lambda/\mu)^k}{k\,!}\right\} p_0 = 1 \; or \; p_0 = \left[\sum_{k=0}^{s} \frac{(\lambda/\mu)^k}{k\,!}\right]^{-1}.$$

Thus $$p_n = \frac{(\lambda/\mu)^n/n\,!}{\sum_{k=0}^{s} \frac{(\lambda/\mu)^k}{k\,!}}\,, \; n = 0, 1, 2, .., s.$$

The probability that an arriving unit is lost to the system (which is the same as that an arrival finds that all the channels are busy and leaves the system or is *lost)* is given by

$$p_n = \frac{(\lambda/\mu)^n/n\,!}{\sum_{k=0}^{s} (\lambda/\mu)^k/k\,!}$$

$$= \frac{a_s/s\,!}{\sum_{k=0}^{s} a^k/k\,!} (putting \; a \; for \; \lambda/\mu)$$

The formula (2.12) is known as *Erlang's loss formula or blocking formula* and is denoted by $B(s, a)$, while (2.11) is known as Erlang's first formula (or simply *Erlang's formula*, the corresponding distribution being truncated Poisson).

Note :

1 Attempts have been made since Erlang's time to generalize Erlang's results. Mention may be made of the works of Pollaczek, Palm, Kosten, Fortet, Sevast' yanov and Takacs. It has been shown that Erlang's formula (2.11) holds for any distribution of service time (having mean $1/\mu$) provided the input is Poisson (with parameter λ), i. e. it holds for the model M/ G/s/s (loss system).

2. See also Example 5(d) Ch. 3.

UNIT XIII NON – MARKOVIAN QUEUING MODELS

13.1 Introduction

So long we have been discussing queuing processes which are either birth and death or non – birth and death processes. They are in either case Markovian and the theory of Markov chain and processes could be applied in their studies. We shall now consider models where the distributions of the inter-arrival time or the service this do not possess the memory-less property, i.e. are not exponential. The process $\{N(t)\}$ giving the state of the system or system size at time t will then be no longer Markovian; however, the analysis of the process can be based on an associated process which is Markovian. Two techniques are generally being used for this purpose. Kendall (1951) use the concept of regeneration point (due to Palm) by suitable choice of regeneration points and extracts, from the process $\{N(t)\}$, Markov chain in discrete time at those points. This is known as the technique of imbedded Markov chains. The second important technique due to Cox(1955) (see also Keilson and Kooharian (1960)) and known as supplementary variable technique, involves inclusion of such variable(s).

We discussed below Kendall's method.

13.2 Queues with Poisson Input: Model M/G/1

Assume that the input process is Poisson with intensity λ and that the service times are i. i. d. random variables having an arbitrary distribution with mean $1/\mu$. Denote the service time by v, its d. f., by $B(t)$, its p. d. f., when it exists, by $b(t)\left(= B'(t)\right)$, and its L.T. by

$$B^*(s) = \int_0^\infty e^{-st} dB(t).$$

Let $t_n, n = 1, 2, .., (t_n = 0)$ be the n^{th} departure epoch, i.e. the instant at which the n^{th} unit completes his service and leaves the system. These points t_n are the regeneration points of the process $\{N(t)\}$. The sequence of points $\{t_n\}$ forms a renewal process. $N(t_n + 0)$, the number in the system immediately after the n^{th} departure has a denumerable state space $\{0,1,2,..\}$, Write $N(t_n+0) \equiv X_n, n = 0,1,..$ and denote by A_n, the random variable giving the number of units that arrive during the service time of the n^{th} unit.

Then $$X_{n+1} = X_n - 1 + A_{n+1}, \qquad if\ X_n \geq 1$$

$$= A_{n+1}, \qquad if\ X_n = 0$$
(3.1)

Now the service times of all the units have the same distribution so that

$$A_n \equiv A\ for\ \ n = 1,2,..We\ have$$

$$\Pr\{\ A = r \mid\ service\ time\ of\ a\ unit\ is\ t\} = \frac{e^{-\lambda t}(\lambda t)^r}{r\,!}$$

And so $$k_r \equiv \Pr\{A = r\} \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^r}{r\,!}\ db(t), \qquad r = 0,1,2,\ldots$$
(3.2)

gives the distribution of A, the number of arrivals during the service time of a unit. The probabilities

$$p_{ij} = Pr\{X_{n+1} = j \mid X_n = i\}$$

Are given by

$$p_{ij} = k_{j-i+1}, \qquad i \geq 1, j \geq i-1$$

$$= 0, \qquad i \geq 1, j < i-1$$
(3.3)

$$p_{oj} \equiv p_{ij} = k_j, \qquad j \geq 0.$$

The relations (3.3) clearly indicate that $\{X_n, n \geq 0\}$ is a Markov chain having t. p. m.

$$P = (p_{ij}) = \begin{bmatrix} k_0 k_1 k_2 & \ldots & \\ k_0 k_1 k_2 & \ldots & \\ 0 & k_1 k_1 & \ldots \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \end{bmatrix}$$
(3.4)

As every state can be reached from every other state, the Markov chain $\{t_n\}$ is irreducible. Again as $p_{ii} \geq 0$, the chain is aperiodic. It can also be shown that, when the traffic intensity $\rho = 1/\mu < 1$, the chain is persistent, non – null and hence ergodic. We can then apply the ergodic theorem of Markov chain (Theorem 2.11).

13. 3. Pollaczek – Khinchine Formula

The limiting probabilities

$$v_j = lim_{n\to\infty} p_{ij}^{(n)}, \quad j = 0, 1, 2, ..$$

Exist and are independent of the initial state i. The probabilities $v = (v_0,\ v_1, \ldots), \sum v_j = 1$,are gives as the unique solutions of

$$V = VP.$$

Let $K(s) = \sum k_j s^j$ *and* $V(s) = \sum v_j s^j$ denote the p. g. f. of the distributions of $\{k_j\}$*and* $\{v_j\}$ respectively.

We have $$K(s) = \sum_{j=0}^{\infty} k_j s^j = \sum_{j=0}^{\infty} s^j \left\{ \int_0^{\infty} \frac{e^{-\lambda t}(\lambda t)^j}{j\,!}\ db(t) \right\}$$

$$= \int_0^{\infty} e^{(-\lambda - \lambda s)\ t}\ dB(t)$$

$$= B^*(\lambda - \lambda s).$$

(3.5)

Hence $$E(A) = K'(1) = -\lambda B^{*(1)}(0) = 1/\mu = \rho.$$
(3.6)

Now $V = VP$ gives an infinite system of equations. Multiplying the $(k+1)st$ equation by $s^k, k = 0, 1, \ldots$ and adding over k, we get, on simplification, for $0 < \rho < 1$,

$$V(s) = \frac{\{1 - K'(1)\}(1 - s)\, K(s)}{K(s) - s}$$ (see Example 8, Ch. 2)

Putting $K'(1) = \rho$ we get

$$V(s) = \frac{(1-\rho)\,(1-s)\,K(s)}{K(s) - s}$$
(3.7)

$$= \frac{(1-\rho)\,(1-s) B^*(\lambda - \lambda s)}{B^*(\lambda - \lambda s) - s}$$

This is known as Pollaczek* - Khinchine (P. K.) formula.

13. 3. 1 Busy Period

The expected duration of the busy period T follows immediately from the result noted in Example 10(a), For $\rho < 1$, we have

$$p_0 = lim_{n\to\infty}\{N(t) = 0\} = \frac{E(I)}{E(I)+E(T)}.$$

Clearly, the idle period I here is exponential with mean $1/\lambda$, since the inter-arrival distribution is so. As stated in the remarks noted earlier in this section, $p_n = v_n$ for all n. From equation (3.7) we find that v_0, the constant term in $V(s)$, is given by $v_0 = 1 - \rho$, so that

$$p_0 = 1 - \rho = \frac{1/\lambda}{1/\lambda+E(T)}$$

Whence $$E(T) = \frac{1}{\mu-\lambda} = \frac{E(v)}{1-\rho},$$ (v being the service time).

(3.8)

Note that $E(T)$ for an M/G/1 queue has the same form as that for an M/M/1 queue.

Thus, given the mean arrival and service rates, the expected duration of a busy period in a queue with Poisson input is independent of the form of the distribution of the service time.

13. 4 GI/M/1 -Model

Here we assume that the service time distribution is exponential with mean $1/\mu$ and that the inter – arrival time is a random variable u, having an arbitrary distribution with mean $1/\lambda$. Denote the d. f. of u by $A(t)$ and its p. d. f., when it exists, by $a(t)$; its L. T. is given by

$$A^*(s) = \int_0^\infty e^{-st}dA(t)$$

(4.1)

So that $$A^{*(k)}(0) \equiv \frac{d^k}{ds^k}A^*(s)\,|_{s=0} = (-1)^k\, E(u^k)$$

(4.2)

And for $k = 1, A^{*(1)}(0) = -1/\lambda$.

Let $t_n, n = 1, 2, \dots, (t_0 = 0)$ be the epoch at which the n^{th} arrival occurs. The process $N(t_n - 0) = Y_n,\ n = 0, 1, 2, \dots$ gives the number in the system immediately before the arrival of the $n^{th}\ unit$. Then

$$Y_{n+1} = Y_n + 1 - B_{n+1},\ if\ \ Y_n \geq 0,\ B_{n+1} \leq Y_{n+1},$$

Where B_{n+1} is the number of units served during $(t_{n+1} - t_n), i.e.$ the inter-arrival time between the n^{th} and $(n+1)st$ unit. As B_n is independent of n, i.e. B_n for all n, its distribution is given by

$$g_r = \Pr\{B = r\} = \int_0^\infty \frac{e^{-\mu r}(\mu t)^r}{r\,!}\, dA(t),\ r = 0, 1, 2, \ldots \quad (4.3)$$

13. 4. 1 Steady State Distribution

The arrival point conditional probabilities

$$p_{ij} = \Pr\{Y_{n+1} = j \mid Y_n = i\}$$

Are given by

$$p_{ij} = g_{i+r-j},\ i+1 \geq j \geq 1,\ i \geq 0 \quad (4.4)$$

$$= 0, i+1 < j,$$

And $\sum_{j=0}^{i+1} p_{ij} = 1.$

Thus $p_{i0} = 1 - \sum_{i=1}^{i+1} p_{ij} = 1 - \sum_{j=1}^{i+1} g_{i+r-j}$

$$= 1 - \sum_{r=0}^{i} g_r = h_i\ (say),\ i \geq 0. \quad (4.5)$$

Since all $p_{ij}{}'s$ depend only on $i\ and\ j$, $\{Y_n,\ n \geq 0\}$ is a Markov chain having t. p. m.

$$P = (p_{ij}) = \begin{bmatrix} h_0 g_0 & 0 & 0 & 0 & \ldots \\ & h_1 g_1 g_0 & 0 & 0 & \ldots \\ & h_2 g_2 g_1 g_0 & 0 & \ldots & \\ \ldots & & \ldots & & \ldots \\ \ldots & & \ldots & & \ldots \end{bmatrix} \quad (4.6)$$

As $g_r > 0$, the chain is irreducible and aperiodic. It can also be shown that it is ergodic, persistent non – null when $\rho < 1$. Thus, when $\rho < 1, i.e.$ in the ergodic case, the limiting arrival point system size probabilities

$$v_j = lim_{n\to\infty} P_{ij}^{(n)}$$

Exist and are given as the unique solution of the system of equations

$$V = VP, \quad (4.7)$$

Where $V = (v_0,\ v_1, ..),\quad \sum v_j = 1.$

We now proceed to find V.

Let $G(s) = \sum_{r=0}^{\infty} g_r s^r$ be the p. g. f. of $\{g_r\}$, $i.e.$ of the r. v. B, the number of units served during an inter – arrival interval. We have

$$G(s) = \sum_{r=0}^{\infty} g_r s^r$$

$$= \sum_{r=0}^{\infty} s^r \int_0^{\infty} \frac{e^{-\mu t}(\mu t)^r}{r\,!}\, dA(t)$$

$$= \int_0^{\infty} e^{-\mu(1-s)t}\, dA(t)$$

$$= A^*\big(\mu(1-s)\big).$$

Also $\quad E(B) = G'(1) - \mu A^{*(1)}\,(0)$

The equations (4.7) can be written as

$$v_0 = \sum_{r=0}^{\infty} v_r h_r \tag{4.8}$$

$$v_j = \sum_{r=0}^{\infty} v_{r+j-1} g_r\,,\ j \geq 1. \tag{4.9}$$

Denoting the displacement operator by $E(so\ E^r(v_k) = v_{k+r}\ etc.)$, we can write (4.9) as a difference equation

$$E\big(v_{j-1}\big) = v_j = \sum_{r=0}^{\infty} g_r E^r\big(v_{j-1}\big),\quad j \geq 1$$

Or, $\quad \{E - \sum_{r=0}^{\infty} g_r E^r\}\big(v_{j-1}\big) = 0$

Or, $\quad \{E - G(E)\}v_{j-1} = 0,\quad j-1 \geq 1.$

This equation can be solved by the method given in Appendix Sec. A. 2. The characteristic equation of the difference equation is

$$r(z) = z - G(z) \equiv z - A^*(\mu - \mu z) = 0.$$

It can be shown that when $G'(1) = 1/\rho > 1, i.e.\ \rho < 1, r(z)$ has only one zero inside $|z| = 1$. Assume that $\rho < 1$, then, if the root of $r(z) = 0\ inside\ |z| = 1$ is denoted by r_0 and the roots on and outside $|z| = 1$ are denoted by $s_0,\ s_1,\ s_2, ...$ then the solution of (4.10) is

$$v_j = C_0 r_0^j + \sum_i D_i s_i^j,\quad j \geq 0.$$

Where $C_0,\ D_i$ are constants.

But since $\sum_i v_j = 1$, $D_i = 0$for all i. Thus, when $\rho < 1$, we get, $v_j = C_0 r_0^j, j \geq 0 \; and \; from$

$\sum_{j=0}^{\infty} v_j = 1$, we get $C_0 = 1 - r_0$, so that

$$v_j = (1 - r_0) r_0^j, j \geq 0.$$

The steady – state arrival point system size has a geometric distribution with mean $r_0/(1 - r_0)$, r_0 being the unique root of $r(z) = 0$ lying inside $|z| = 1$.

Remark:

We take as regeneration points the arrival epochs in case of GI/M/1 model and departure epochs in case of M/G/1 model. The queuing processes occurring in the two models M/G/1 and GI/M/1 are non Markovian; however by extracting processes at these regeneration points, it has been possible to obtain Markov chains from those processes. The embedded Markov chains indicate *system stateat regeneration* points. What one needs also is the distribution of *general time system state*.

Let us now consider the queuing process of GI/M/1, having for its embedded Markov chain $\{Y_n\}$, where Y_n in the system size immediately prior of n^{yh} arrival. We have obtained $\{v_j\}$, the limit distribution of $\{Y_n\}$. Define $Z(t) = Y_n, t_n \leq t \leq t_{n+1}$. Then $\{Z(t)\}, where \; Z(t)$ gives the system size at the most recent arrival, is a semi – Markov process having $\{Y_n\}$ for its embedded Markov chain. $\{Y_n, \; t_n\}$ is an irreducible Markov renewal process.

Let $N(t)$ denote the system size at an arbitrary time t. Denote

$$f_{ij}(t) = \Pr\{Z(t) = j \mid Z(0) = i\}$$

$$p_{ij}(t) = \Pr\{N(t) = j \mid Z(0) = i\};$$

The $f_j = lim_{t \to \infty} f_{ij}(t)$gives the limiting probability that the system size of the $s - M.P. Z(t) \; is \; j$, whereas $p_j = \lim \; p_{ij}(t) \; as \; \to \infty$, gives the limiting probability that the system size of the general time process $N(t) \; is \; j$.

We have to look for relationship, if any, existing between the three limit distributions $\{v_j\}, , \{f_j\}$and $\{p_j\}$.

From theorem 7.1, we find that $v_j \; and \; f_j$ are related as follows:

$$f_j = \frac{v_j M_j}{\sum_i v_i M_i},$$ (4.12)

Where M_i is the expected time spent in the state I during each visit.

The transitions, in case of this model, occur at the arrival points which are its regeneration points. Thus M_i, the expected time spent in a state I during each visit is the expected inter-arrival time. IN other words, $M_i = 1/\lambda$ for all i. Putting the value of M_i in (4.13), we get

$$f_j = \frac{v_j\ /\lambda}{\sum_i v_i\ /\lambda} = \frac{v_j}{\sum_i v_i} = v_j \quad for\ all\ j,$$

$$= (1 - r_0)r_0^j \text{ (from (4.11)).}$$

In order to obtain p_j in terms of $f_j(and\ terms\ of\ v_j)$we have to make use of the relationship between $f_j\ and\ p_j$.

It is shown that for this particular model,

$$p_j = (\lambda/\mu)\ f_j/r_0,$$

$$= (\lambda/\mu)(1 - r_0)r_0^{j-1} = \rho v_{j-1},\ j \geq 1$$

And $$p_0 = 1 - \lambda/\mu.$$

We get

Mean $= \rho/(1 - r_0)$ and variance $= \rho(1 - \rho + r_0)/(1 - r_0)^2$.

UNIT – XIV Non -Birth -Death Queuing Processes :Bulk Queues:

14.1 THE $M^X/G/1$ QUEUE:

A useful model is the single – server $M^X/G/1$queue where batches of customers arrive according to a Poisson process with rate λ and the batch size X has a discrete probability distribution $\{\beta_j,\ j = 1,\ 2, ..\}$ with finite mean β. The customers are served individually by a single server. The service times of the customers are independent random variables with a common probability distribution function $B(t)$. Denoting by the random variable S the service time of a customer, it is assumed that the server utilization ρ defined by

$$\rho = \lambda \beta E(s)$$

Is smaller than 1. The analysis for the $M/G/1$ queue can be extended to the $M^X/G/1$ **queue.** In section 9. 3. 1 we give an algorithm for the state probabilities. The computation of the waiting – time probabilities is discussed in section 9. 3. 2.

14. 2 The State Probabilities

The stochastic process $\{L(t),\ t \geq 0\}$ describing the number of customers in the system is regenerative. The process regenerates itself each time an arriving batch finds the system empty. The cycle length has a continuous distribution with finite mean. Thus the process $\{L(t)\}$ has a limiting distribution $\{p_j\}$. The probability p_j can be interpreted as the long – run fraction of time that j customers are in the system. The probability p_0 allows for the explicit expression

$$p_0 = 1 - \rho.$$

To see this, we apply the ‘ reward principle’ that was used in section 2.3 to obtain Little’s formula. Assume that the system earn a reward at rate 1 whenever a customer is in service. Then the average reward per time unit represents the fraction of time that the server is busy. The long – run average reward earned per customer is equal to $E(s)$, while the long – run average arrival rate of customers is $\lambda\beta$. Hence the long – run average reward earned per time unit equals $\lambda\beta\ E(s)$. The long – run fraction of time that the server is busy equals $1 - p_0$. This shows that $1 - p_0 = \lambda\beta\ E(s) = \rho$.A recursion scheme for thep_j is given in the following theorem.

Theorem 14. 3

The state probabilities p_j satisfy the recursion

$$p_j = \lambda p_0 \sum_{s=1}^{j} \beta_s a_{j-s} + \lambda \sum_{k=1}^{j} \left(\sum_{i=0}^{k} p_i \sum_{s>k-i} \beta_s\right) a_{j-s}, \quad j - 1, 2, \ldots,$$

Where

$$a_n = \int_0^{\infty} r_n(t)\{1 - B(t)\}dt, \quad n = 0,\ 1, \ldots$$

with $r_n(t) = P\{$ *a total of n customers will arrive in (0, t)* $\}$.

Proof: The proof is along the same lines as the proof of Theorem 9. 2. 1. The only modification is with respect to the up – and down-crossing relation (9. 2. 1). We now use the following up – and down – crossing argument : the number of down – crossings from a state in the set $\{k + 1, k + 2, \ldots\}$ to a state outside this during one cycle equals the number of up-crossings from a state outside the set $\{k + 1, k + 2, \ldots\}$ to a state in this set during one cycle. Thus relation (9.2.5) generalize to

$$E(N_k) = \sum_{i=0}^{k} E(T_i)\lambda \sum_{s>k-i} \beta_s, \quad k = 0, 1, \ldots$$

The remainder of the proof is analogous to the proof of Theorem 9. 2. 1.

The recursion scheme (9. 3. 2) is not as easy to apply as the recursion scheme (9. 2. 1). The reason is that the computation of the constants a_n is quite burden -some. In general, numerical integration must be used, where each function evaluation in the integration procedure requires an application of Adelson’s recursion scheme for the computation of the compound Poisson probabilities $r_n(t), n \geq 0$;

The best general – purpose approach for the computation of the state probabilities is the discrete FFT method. An explicit expression for the generating function

$$P(z) = \sum_{j=0}^{\infty} p_j z^j, \quad |z| \leq 1$$

Can be given. It is a matter of tedious algebra to derive from ((9. 3. 2)) that

$$P(z) = (1-\rho)\frac{1-\lambda\alpha(z)\ \{1-G(z\})}{1-\lambda\alpha(z)\ \{1-G(z\})/(1-z)},$$
(9. 3. 3)

Where

$$G(z) = \sum_{j=0}^{\infty} \beta_j z^j \ and\ \alpha(z) = \int_0^{\infty} e^{-\lambda\{1-G(z)t\}}\,(1-B(t))dt.$$

The derivation uses that $e^{-\lambda\{1-G(z)t\}}$ is the generating function of the compound Poisson probabilities $r_n(t)$; see Theorem 1. 2.1. Moreover, the derivation uses that the generating function of the convolution of two discrete probability distributions is the product of the generating functions of the two probabilities distributions. The other details of the derivation of (9. 3. 3) are left to the reader. For constant and phase-type services, no numerical integration is required to evaluate the function $\alpha(z)$ in the discrete FFT method.

Asymptotic expansion:

The state probabilities allow for an asymptotic expansion when it is assumed that the batch-size distribution and the service-time distribution are not heavy-tailed. Let us make the following assumption.

Assumption 14. 3. 1

(a) The convergence radius$R$ of $G(z) = \sum_{j=1}^{\infty} \beta_j z^j$ is larger than 1. Moreover, $\int_0^{\infty} e^{st}\{1-B(t)\}dt < \infty$for some $s > 0$.

(b) $lim_{s\to B} \int_0^{\infty} e^{st}\{1-B(t)\}dt = \infty$, where B is the supremum over all s with

$$\int_0^{\infty} e^{st}\{1-B(t)\}dt] < \infty.$$

(c) $lim_{x\to R_0} G(x) = 1 + B/\lambda$ for some number $R_0\ with\ 1 < R_0 \leq R$.

Under this assumption we obtain from Theorem C. 1 in Appendix C that

$$p_j \sim \sigma\, \tau^j \; as\; j \to \infty,$$
(9. 3.4)

Where τ is the unique solution to the equation

$$\lambda\alpha(\tau)\{1 - G(\tau)\} = 1 - \tau$$
(9. 3. 5)

On $(1,\ R_0)$ and the constant σ is given by

$$\sigma = (1-\rho)(1-\tau)\left[\lambda\alpha'(\tau)\{1 - G(\tau)\} - \frac{(1-\tau)G'(\tau)}{1-G(\tau)} + 1\right]^{-1}.$$
(9. 3. 6)

A formula for the average queue size

The long – run average number of customers in queue is $L_q = \sum_{j=1}^{\infty}(j-1)p_j$. Using the relation $P'(1) = \sum_{j=1}^{\infty} jp_j$, we obtain after some algebra from (9. 3. 3) that

$$L_q = \frac{1}{2}(1 + c_S^2)\frac{\rho^2}{1-\rho} + \frac{\rho}{2(1-\rho)}\left[\frac{E(X^2)}{E(X)} - 1\right],$$

Where X denotes the batch size. Note that the first part the expression for L_q gives the average size in the standard M/G/1 queue, while the second part reflects the additional effect of the batch size. The formula for L_q implies directly a formula for the long – run average delay in queue per customer. By Little's formula $L_q = \lambda\beta W_q$.

14. 3. The Waiting – Time Probabilities

The concept of waiting distribution is more subtle for the case of batch arrivals than for the case of single arrivals. Let us assume that customers from each arrival group are numbered as 1, 2,… Service to customers from the same arrival group is given in the order in which those customers are numbered. For customers from different batches the service is in order of arrival. Define the random variable D_n as the delay in queue of the customer who receives the n^{th} service. In the batch – arrival queue, $lim_{n\to\infty} P\{D_n \le x\}$ need not exist. To see this, consider the particular case of a constant batch size of 2. Then $P\{D_n > 0\} = 1$ for n even and $P\{D_n > 0\} < 1$ for n odd. The limit

$$W_q(x) = lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} P\{D_k \le x\}, \qquad x \ge 0$$

Always exists. To see this, fix x and imagine that a reward of 1 is earned for each customer whose delay in queue is no more than x. Using renewal – reward theory, it can be shown that the limit $W_q(x)$ exists and represents the long – run fraction of customers whose delay in queue is no more than x. If the batch size distribution is non – arithmetic, then $lim_{n\to\infty} P\{D_n \le x\}$exists and equals $W_q(x)$.

Denote by

$$b^*(s) = \int_0^\infty e^{-sx} b(x)dx$$

The Laplace transform of the probability density $b(x)$ of the service time of a customer. Let $\beta^*_{SC}(s)$ be the Laplace transform of the probability density of the total time needed to serve all customers from one batch. It is left to the reader to verify that

$$\beta^*_{SC}(s) = \sum_{k=1}^\infty \beta_k [b^*(s)]^k = G\big(b^*(s)\big).$$

The following result now holds:

$$\int_0^\infty e^{-sx}\{1 - W_q(x)\}dx = \frac{1 - W^*_{SC}(s)W^*_r(s)}{s}, \quad (9.\ 3.\ 7)$$

Where

$$W_{SC}(s^*) = \frac{(1-\rho)s}{s-\lambda+\lambda\beta^*_{SC}(s)} \quad and \quad W^*_r(s) = \frac{1-G(b^*(s))}{\beta[1-b^*(s)]}$$

With $\beta = \sum_{k=1}^\infty k\beta_k$ denoting the average batch size. The waiting – time probabilities $W_q(x)$ can be numerically obtained from (9. 3. 7) by using numerical Laplace inversion .

We give only a heuristic sketch of the oroof of (9. 3. 7). A rigorous treatment is given in Van Ommeren (1988). An essential part of the proof is the following result. For $k = 1, 2, \ldots$, let

$\eta k =$ the long – run fraction of customers taking the k^{th} position in their batch.

Then it holds that

$$\eta k = \frac{1}{\beta}\sum_{j=k}^\infty \beta_j, \qquad k = 1, 2, \ldots. \quad (9.\ 3.\ 8)$$

To prove this result, fix k and imagine that a reward of 1 is earned for each customer taking the k^{th} position in its batch. Then the long – run average reward per customer is ηk by definition. By the renewal – reward theorem, the long –run average reward per customer equals the expected reward $\sum_{j=k}^{\infty} \beta_j$earned for a single batch divided by the expected batch size β. This gives (9. 3. 8). Consider now a test customer belonging to a batch that arrives when the system has reached steady state. Denote by $D^{(\infty)}$ the delay in queue of this test customer. The delay $D^{(\infty)}$ can be written as $D^{(\infty)} = X_0 + X_1$, where X_0 is the delay caused by the customers present just before the batch of the test customer arrives and X_1 is the delay caused by customers belonging to the batch of the test customer. The random variables X_0 and X_1 are independent of each other and so $E\left(e^{-sD^{(\infty)}}\right) = E(e^{-sX_0})E(e^{-sX_1})$, Assuming that the position of the test customer in the batch is distributed according to $\{\eta k\}$, we have by (9. 3. 8) that

$$E(e^{-sX_1}) = \sum_{k=1}^{\infty} \eta k\, [b^*(s)]^{k-1} = \frac{1}{\beta} \sum_{k=1}^{\infty} [b^*(s)]^{k-1} \sum_{j=k}^{\infty} \beta_j$$

$$= \frac{1}{\beta} \sum_{j=k}^{\infty} \beta_j \sum_{k=1}^{\infty} [b^*(s)]^{k-1} = \frac{1-G(b^*(s))}{\beta[1-b^*(s)]} .$$

To find $E(e^{-sX_0})$, note that an arriving group of customers can be considered as a singly arriving super-customer. The probability density of the total time. The probability density of the total time to serve a super – customer has the Laplace transform $\beta_{SC}^*(s)$. In other words, the delay in queue of the first customer of each batch can be described by a standard M/G/1 queue for which the service – time density has the Laplace transform $\beta_{SC}^*(s)$. Thus, using the result for the M/G/1 queue.

$$E(e^{-sX_0}) = \frac{(1-\rho)s}{s-\lambda+\lambda\beta_{SC}^*(s)} .$$

Since $\int_0^{\infty} e^{-sx}\{1 - W_q(x)\}dx = s^{-1}\left[1 - E\left(e^{-sD^{(\infty)}}\right)\right]$ by relation in Appendix E, we have now derived (9. 3. 7) heuristically.

14.4.Alternative algorithm:

A simpler algorithm than numerical Laplace inversion can be given for the $M^X/D/1$ queue with deterministic services, This alternative algorithm is discussed in Section 9. 5. 3 in the more general context of the $M^X/D/1$ queue. A simple algorithm is also possible when the service time of a customer is a mixture of Erlangian distributions with the same scale parameters. In this case the service time of a customer can be interpreted as a random sum of independent phases each having an exponentially distributed length with the same mean. The $M^X/G/1$ queue with generalized Erlangian services is in fact an $M^Y/M/1$ queue in which the batch size Y is distributed as the total number of service phases generated by all customers in one batch. For this particular $M^X/G/1$ queue the waiting - time probabilities $W_q(x)$ can be computed by a modification of the algorithm gives in Example 5. 5. 1.

Approximations for the waiting – time probabilities

Suppose that Assumption 9. 3. 1 is satisfied and let $b(t)$ denote the density of the service – time distribution function $B(t)$. Then the following asymptotic expansion applies:

$$\mathbf{1} - W_q(x) \sim \gamma e^{-\delta x}\ as\ x \to \infty,$$

Where δ is the smallest positive solution to

$$\sum_{j=1}^{\infty} \beta_j \left\{\int_0^{\infty} e^{\delta t}\ b(t)dt\right\}^j = 1 + \frac{\delta}{\lambda}$$

And the constant γ is given by

$$\gamma = \frac{(1-\rho)\delta}{\lambda\beta}\left[1 - \lambda \int_0^{\infty} e^{\delta t}\ b(t)dt\ \sum_{j=1}^{\infty} j\beta_j \left\{\int_0^{\infty} e^{\delta t}\ b(t)dt\right\}^{j-1}\right]^{-1}$$

$$\boldsymbol{X}\left[\mathbf{1} - \int_0^{\infty} e^{\delta t}\ b(t)dt\right]^{-\mathbf{1}}.$$

14.5 The $M^X/G/c$ Queue

In the $M^X/G/c$ queue the customers arrive in batches rather than singly. The arrival process of batches is a Poisson process with rate λ. The batch size has a probability distribution $\{\beta_j,\ j = 1, 2, ..\}$ with finite mean β. The service times of the customers are independent of each other and have a general distribution with mean $E(s)$. There are c identical servers. It is assumed that the server utilization ρ, defined by

$$\rho = \frac{\lambda\beta E(s)}{c},$$

Is smaller than 1. The customers from different batches are served in order of arrival and customers from the same batch are served in the same order as their positions in the batch. A computationally tractable analysis can only be given for the special cases of exponential services and deterministic services. We first analyse these two special cases. Next we discuss a two – moment approximation for the general $M^X/G/c$ queue.

14.5.1 The $M^X/M/c$ queue

The process $\{L(t)\}$ describing the number of customers present is a continuous – time Markov chain. Equating the rate at which the process leaves the set of states $\{i, i+1, ..\}$ to the rate at which process this set of states, we find for the state probabilities p_j the recursion scheme

$$\min(i, c)\,\mu p_i = \sum_{k=0}^{i-1} p_k \lambda \;\; \sum_{s \geq i-k} \beta_s, \;\; i = 1, 2, ... \qquad (9.\ 6.\ 32)$$

Where $\mu = \frac{1}{E(S)}$, Starting with $\overline{p_0} := 1$, we successively compute $\overline{p_1}, \overline{p_2}, ..$ and next obtain desired p_i by normalization. The normalization can be based on Little's relation

$$\sum_{j=0}^{c-1} j p_j + c\left(1 - \sum_{j=0}^{c-1} p_j\right) = c\rho \qquad (9.\ 6.33)$$

Stating that the average number of busy servers equals $c\rho$. The computational effort of the recursion scheme can be reduced by using the asymptotic expansion,

$$p_j \sim \sigma \tau^{-j} \; as \; j \; \rightarrow \infty, \qquad (9.\ 6.\ 34)$$

Where τ is the unique solution of the equation

$$\lambda \tau [1 - \beta(\tau)] = c\mu(1 - \tau) \qquad (9.\ 6.\ 35)$$

On the interval $(1, R)$ and the constant σ is given by

$$\sigma = \frac{(\tau - 1)\sum_{i=0}^{c-1}(c - i) p_i \tau^i / c}{1 - \lambda \tau^2 \beta'(\tau)/(c\mu)}. \qquad (9.\ 6.\ 36)$$

Here $\beta(z) = \sum_{j=1}^{\infty} \beta_j z^j$ and R is the convergence radius of the power series $\beta(z)$. To establish the asymptotic expansion, it is assumed that $R > 1$. In other words, the batch – size distribution is not heavy – tailed. The derivation of the asymptotic expansion (9. 6. 34) is routine. Define the generating function $P(z) = \sum_{j=0}^{\infty} p_j z^j, |z| \leq 1$.It is a matter of simple algebra to derive from (9. 6. 32) that

$$P(z) = \frac{(1/c)\sum_{i=0}^{c-1}(c - i) p_i z^i}{1 - \lambda z\, \{1 - \beta(z)\}/\{c\mu(1 - z)\}}.$$

Next, by applying Theorem C.1 in Appendix C, we obtain (9. 6. 34).

From the generating function we also derive after considerable algebra that the long – run average queue size is given by

$$L_q = \frac{1}{c(1-\rho)} \sum_{i=0}^{c-1} j(c-i)p_j + \frac{\rho}{2(1-\rho)} \left\{ \frac{E(X^2)}{E(X)} - 1 \right\} + \frac{\rho}{1-\rho} - c\rho,$$

Where the random variable X denotes the batch size.

Next we discuss the computation of the steady – state probability distribution function $W_q(x)$ of the waiting time of a customer. The function $W_q(x)$, is defined in the same way as in section 9. 3. 2. To find $W_q(x)$, we need the probabilities

$z_j =$ the long – run fraction of customers who have j other customers in front of them just after arrival , $j = 0, 1, \ldots$

The delay in queue of a customer who has $j \geq c$ other customers in front of him just after arrival is the sum of $j - c + 1$independent exponentials with common mean $1/c\mu$. Hence this conditional waiting time has an E_{j-c+1} distribution and so

$$1 - W_q(x) = \sum_{j=c}^{\infty} z_j \sum_{k=0}^{j-c} e^{-c\mu x} \frac{(c\mu x)^k}{k\,!}, \quad x \geq 0.$$

A computationally better representation for $W_q(x)$ is

$$1 - W_q(x) = \sum_{k=0}^{\infty} e^{-c\mu x} \frac{(c\mu x)^k}{k\,!} \left(1 - \sum_{j=0}^{k+c-1} z_j\right), \quad x \geq 0\,.$$
(9. 6. 37)

The probabilities z_j are easily expressed in terms of the p_j. To do so, let

$$\eta k = \frac{1}{\beta} \sum_{j=k}^{\infty} \beta_j, \quad k = 1, 2, \ldots$$

Then, as shown in Section 9. 3. 2, The probability ηk gives the long – run fraction of customers who take the k^{th} position in their batch. Since the long – run fraction of batches finding m other customers present upon arrival equals p_m, we find

$$z_j = \sum_{m=0}^{j} p_m \eta_{j-m+1}, \quad j = 0, 1, \ldots$$

For the case of exponential services this formula can be considerably simplified. Using the recursion relation (9. 6. 32), we have

$$z_j = \frac{\mu}{\lambda\beta} \min(j+1, c) p_{j+1}, \quad j = 0,\ 1, \ldots.$$
(9. 6. 38)

This completes the specification of the exact algorithm (9. 6. 37) for the computation of the waiting – time probabilities $W_q(x)$. The computational effort can further be reduced by using an asymptotic expansion for $1 - W_q(x)$. Inserting (9. 6. 34) and (9. 6. 38) into (9. 6. 37), we find after some algebra that

$$1 - W_q(x) \sim \frac{\sigma\tau^{-c}}{\tau - 1} e^{-c\mu\left(1-\frac{1}{\tau}\right)x} \quad as\ x \to \infty,$$
(9. 6. 39)

Where τ *and* σ are given by (9. 6. 35) and (9. 6. 36).

14.6 The $M^X/D/c$ queue

Suppose that the service time of each customer is a constant D. Denoting by $p_j(t)$ the probability that j customers are present at time t, we find by the same arguments as used in Section 9. 6. 2 that

$$p_j(t + D) = \sum_{k=0}^{c} p_k(t)\, r_j(D) + \sum_{k=c+1}^{c+j} p_k(t)\ r_{j-k+c}(D), \qquad j = 0, 1, \dots$$

Where the compound Poisson probability $r_j(D)$ is defined by

$r_j(D) =$ the probability that exactly j customers arrive during a given time interval of length D, $j = 0, 1, \dots$

Letting $t \to \infty$, we find the system of linear equations

$$p_j = r_j(D) \sum_{k=0}^{c} p_k + \sum_{k=c+1}^{c+j} r_{j-k+c}(D) p_k, \qquad j = 0, 1, \dots$$
(9. 6. 40)

Together with the normalizing equation $\sum_{j=0}^{\infty} p_j = 1$. Just as in the M/D/c case, this infinite system of equations by using the geometric tail behavior of the p_j. It holds that

$$p_j \sim \sigma\tau^{-j} \quad as \quad j \to \infty,$$
(9. 6. 41)

Where τ is the unique root of the equation

$$\tau^c e^{\lambda D\{1-\beta(\tau)\}} = 1$$
(9. 6. 42)

On the interval (1, R) and the constant σ is given by

$$\sigma = [c - \lambda D\tau\beta'(\tau)]^{-1} \sum_{j=0}^{c-1} p_j(\tau^j - \tau^c).$$
(9. 6. 43)

As before ,$\beta(z) = \sum_{j=1}^{\infty} \beta_j z^j$ and the number R denotes the convergence radius of power series $\beta(z)$.

It is assumed that $R > 1$.

In general however , it is computationally simpler to compute the state probabilities p_j by applying the discrete FFT method to the generating function $P(z) = \sum_{j=1}^{\infty} p_j z^j$. In the same way as (9. 6. 6) we derived, we obtain

$$P(z) = \frac{\sum_{j=0}^{c-1} p_j(z^j - z^c).}{1 - z^c e^{\lambda D\{1-\beta(z)\}}},$$
(9. 6. 44)

Since the generating function of the compound Poisson probabilities $r_j(D)$ is given by $e^{\lambda D\{1-\beta(z)\}}$; Before the discrete FFT method can be applied, the unknown probabilities $p_0, \dots.. p_{c-1}$ must be removed from (9. 6. 44). To do so, we proceed in the same way as in Section 9. 6. 1 and rewrite $P(z)$ in the explicit form

$$P(z) = \frac{c(1-\rho)(1-z)}{1 - z^c e^{\lambda D\{1-\beta(z)\}}} \prod_{k=1}^{c-1} \left(\frac{z - zk}{1 - zk}\right),$$
(9. 6. 45)

Where $z_0 = 1,\ z_1, \dots, z_{c-1}$ are the c distinct roots of $z^c e^{\lambda D\{1-\beta(z)\}} = 1$ inside or on the unit circle. The computation of the roots $z_1, \dots. z_{c-1}$ is discussed in Appendix G. The asymptotic expansion (9. 6. 41) follows from the generating function (9. 6. 44) and Theorem C. 1 in Appendix C. Also, we obtain after considerable algebra from (9. 6. 44) that the long – run average size is given by

$$L_q = \frac{1}{2c(1-\rho)}\Big[(c\rho)^2 - c(c-1) + \sum_{j=2}^{c-2}\{c(c-1) - j(j-1)\}p_j +$$
$c\rho EX2EX-1,$

Where the random variable X denotes the batch size. This relation can be used as an accuracy check on the calculated values of the probabilities p_j.

Waiting – time probabilities in the $M^X/D/c$ queue

In the batch – arrival $M^X/D/c$ queue, the waiting – time probability $W_q(x)$ is defined as the long – run fraction of customers whose time in queue is no more than $x,\ x \geq 0$, The expression (9. 6. 9) for $W_q(x)$ in the M/D/c queue can be extended to the $\boldsymbol{M^X/G/c}$queue. For any x with $(k-1)\,D \leq x < kD \; and \;\; k=1,2,\ldots$, it holds that

$$W_q(x) = \sum_{m=0}^{kc-1} \eta_{m+1} \sum_{j=0}^{kc-1-m} Q_{kc-1-m-j} r_j(kD-x) \qquad (9.\,6.\,46)$$

Where $Q_j = \sum_{m=0}^{kc-1} p_i \; for \;\; j=0,1,\ldots$ and the probability η_r is defined by

$$\eta_r = \frac{1}{\beta}\sum_{j=r}^{\infty}\beta_j\,, \qquad r=1,2,\ldots.$$

This result is due to Franx (2002). Its proof will be omitted . The asymptotic expansion

$$1 - W_q(x) \sim \gamma\, e^{-\lambda[\beta(\tau)-1]x} \quad as\; x \to \infty \qquad (9.\,6.\,47)$$

Holds with

$$\gamma = \frac{\sigma[\beta(\tau)-1]}{(\tau-1)^2\,\tau^{c-1}\,\beta},$$

Where $\tau \; and \; \sigma$ are given by (9. 6. 42) and (9. 6. 43). This result can be derived in a similar way as expansion (9. 6. 11) for the M/D/c queue was obtained.

14.6 The$\boldsymbol{M^X/D/c}$ queue

An exact and tractable solution for the $M^X/D/c$ queue is in general not possible except for the special cases of deterministic and exponential services. Using the solutions for these special cases, we can useful approximations for the general $\boldsymbol{M^X/G/c}$ **queue. A practically useful approximation to the average delay in queue per customer is**

$$W_q^{app} = (1-c_s^2)W_q(det) + c_s^2 W_q(exp),$$

Provided thatc_s^2 is not too large $(say, 0 \leq \; c_s^2 \leq 2)$ and the traffic load is not very small. It was pointed out in Section 9. 3 that the first – order approximation $\frac{1}{2}(1+c_s^2)W_q(exp)$ is not applicable in the batch – arrival queue. A two – moment approximation to the percentiles $\eta(P)$ of the waiting – time distribution of the delayed customers is provided by

$$\eta_{app}(p) = (1-c_s^2)\eta_{det}(p) + c_s^2\eta_{exp}(p), \qquad 0 < p < 1.$$

However, it turns out that in the batch – arrival case the two – moment approximation to $\eta(p)$ works only for the higher percentiles. Fortunately, higher percentiles are usually the percentiles of interest in practice. Table 14. 6. 3 gives the $M^X/E_2/c$ queue the exact and approximate values of the conditional waiting – time percentiles $\eta(P)$ both for the case of a constant batch size and the case of a geometrically distribution batch size. In both cases the mean batch size $E(X) = 3$. The normalization $E(S) = 1$is used for the service time. The percentiles $\eta_{exp}(p)$ for exponential service and $\eta_{det}(p)$ for deterministic services have been computed from the asymptotic expansion (9. 6. 39) and (9. 6. 47). These asymptotic expansions already apply for moderate measure for the traffic load is the probability that all servers are simultaneously busy. This probability is by $P_B = 1 - \sum_{j=0}^{c-1} p_j$. As a rule of thumb, the asymptotic expansions can be for practical purpose for $x \geq \frac{E(X)E(S)}{\sqrt{c}}$ $when\ P_B \geq 0.2$.

BEST WISHES____________________________________

References:

1. Cinlar, E., Introduction to Stochastic Processes , Prentice Hall Inc., Englewood Cliffs New Jersey.
2. Bartlett, M.S., Stochastic processes, 3rd edition, CUP, Cambridge (1978)
3. Doob, J. L. Stochastic Processes, Wiley, New York.(1953)
4. Trivedi, K.S., Probability and Statisticsin Engineering Systems , Prentice Hall International(2002)
5. Chung, Kai.Lai., Markov Chains with Stationary Transition Probabilities,
6. Cox, D.R. and Miller ,H.D., The theory of stochastic Processes , Methuen, London(1965)
7. Tijims, H.C. A First Course in Stochastic Models, Wiley , New York, (2003).
8. Karlin, S. and H. M. Taylor, A First course in Stochastic Processess, 2nd edition Academic Press, New York.(1975).
9. Feller, W. , An introduction to stochastic process and its applications.V-1, 3rd edition Wiley , New York. (1966)
10. Gross,D. and Harris ,C.M. Fundamentals of Queuing theory, 2nd edition , Wiley New York.