UNIT 4 BRANCHING PROCESSES

Structure **Page No**

4.1 INTRODUCTION

Developments in this area of study in applied stochastic processes were initiated around 1874 through the research of Francis Galton and H. W. Watson in respect of the extinction of family surnames.

As is the practice in almost all societies, the family surname runs down the generations through the male offspring of the family. It is possible that all the offspring of a particular generation fail to reproduce a single child. Obviously then, in the subsequent generation there will be nobody to carry forward the family name, which then becomes extinct.

Since the number of children to be born to a couple cannot be exactly predicted because of inherent randomness, the obvious issue of interest would concern the probabilities of extinction of the family name at subsequent generations, or the probability of extinction. In a family tree, the branches are the descriptions of the generations hence the name of the process. We shall begin the unit by discussing the branching process and its mathematical formulations in terms of generating function, expectations and variance in Sec. 4.2. In Sec. 4.3, we will continue our discussion to the probability of extinction which is followed by some of the limiting properties of the probability of extinction.

Obviously, branching processes have considerable relevance in the context of studies in physics (neutrons), biology (bacteria), polymer chemistry (bonds), etc.

Objectives

After reading this unit, you should be able to:

- describe the branching process with examples;
- obtain the probability generating function, expectations and variance for a branching process;
- define the probabilities of extinction;
- express the various limiting properties for the probability of extinction.

4.2 MATHEMATICAL FORMULATIONS

Consider a population of X_0 elements, biological or otherwise (living beings or organisms like cells, neutrons, etc.), capable of reproducing offspring of the same kind. Suppose that each element, by the end of its life time, produces j offspring with probability p_j, $j = 0, 1, 2, \ldots,$ independently of any other offspring. Here, we are interested in the size of successive generations. We note that X_0 is the size of the 0-th generation, and these elements of this initial set are called **ancestors**. All the offspring born to X_0 elements of generation- 0 comprise the first generation. The elements generated by the elements of 0-generation are called the **direct descendents** of the ancestors. Let X_1 denote their total number, so that the size of the first generation is X_1 (which is, obviously, a random number).

Again, each element of the first generation (X_1 in number) would produce offspring according to the probability law $\{p_j\}$ independently of others, to form second generation and the process of reproduction continues in this manner.

Let X_n denotes the size of the n-th generation where $n = 0, 1, 2, 3, \ldots$.

Mathematically, in view of the above structure,

$$X_{n+1} = \xi_1 + \xi_2 + \cdots + \xi_{x_n},\ n = 0, 1, 2, \ldots$$

$$p_{ij} = P(X_{n+1} = j \mid X_n = i) = P(\xi_1 + \xi_2 + \cdots + \xi_i = j);\ i = 1, 2, \ldots, j = 0, 1, 2, \ldots \quad (1)$$

where $\xi_1, \xi_2, \ldots, \xi_i$, etc. are independent and identically distributed **(iid)** random variables with $P(\xi_k = r) = p_r$, where $r = 0, 1, 2, 3, \ldots$, and $k = 1, 2, \ldots, i$ and $p_r \geq 0, \sum_{n=0}^{\infty} p_r = 1$.

The sequence $\{X_n\}$ constitutes a branching process which is a special case of a Markov chain with state space $(0, 1, 2, \ldots)$.

Now let us see the real life situations of the branching process defined above.

(i) Suppose, a series of plates are set up along the path of a current of electrons emitted by a mechanism. After hitting a reflector, an electron splits into a random number of new electrons. Suppose, X_0 is the number of electrons that hit the first plate and split into X_1 electrons which, in turn, after hitting the second plate split into X_2 electrons, and so on. The sequence $\{X_n, n \geq 0\}$ is a branching process.

(ii) Biological entities, such as human beings, and animals, that reproduce are examples of branching models.

(iii) Consider that a person with an infectious disease may transmit this disease to others. Now the number of infected persons in the n-th generation forms a branching process where all the infected persons of a generation are the offsprings of the immediately preceding generation.

Obviously, we can see that $\{X_n\}$ is a Markov chain with a transition probability matrix $((p_{ij}))$. This stochastic process is called a **discrete time branching process** or a **Galton-Watson process**. It is defined to be sub-critical, critical or super-critical as $\mu = \sum_{j=0}^{\infty} jp_j$ is '<', '=' or '>' 1. Notice that μ is the average number offspring of an individual element. For algebraic simplicity, let us take $X_0 = 1$. Then the probability generating function **(p.g.f.)** of the offspring distribution $\{p_j\}$ is given by

$$\phi(s) = Es^{X_1} = \sum_{j=0}^{\infty} s^j p_j,\quad 0 \leq s \leq 1 \quad (2)$$

We denote the p.g.f. of X_n by $\phi_n(s)$ so that $\phi_n(s) = Es^{X_n},\ 0 \leq s \leq 1$.
Notice that for $n \geq 1$,

$$\phi_{n+1}(s) = Es^{X_{n+1}}$$
$$= \sum_{j=0}^{\infty} s^j P(X_{n+1} = j)$$
$$= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} P(X_{n+1} = j \mid X_n = i)\, P(X_n = i) s^j$$
$$= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} P(\xi_1 + \xi_2 + \cdots + \xi_i = j)\, P(X_n = i) s^j,$$

[Where, naturally, $P(X_{n+1} = j \mid X_n = 0) = 0$ for all $j > 0$, as there is none in generation n to produce offspring].

$$= \sum_{i=0}^{\infty} P(X_n = i) \sum_{j=0}^{\infty} P(\xi_1 + \xi_2 + \cdots + \xi_i = j)\, s^j$$
$$= \sum_{i=0}^{\infty} P(X_n = i)\, \{\phi\,(s)\}^i = \phi_n(\phi(s)), \text{ for } n = 1, 2, \ldots \text{ [using Eqn.(2)]} \qquad (3)$$

since, $\xi_1, \xi_2, \ldots, \xi_i$ being independent, the p.g.f. of their sum is the product of their p.g.f's which happen to be $\phi(s)$ for each ξ_r, $r = 1, 2, \ldots, i$. Here, it may be noted that $\phi_1(s) = \phi(s)$, is the generating functions of X_1.

By induction then,

$$\phi_{n+1}(s) = \phi_n(\phi(s))$$
$$= \phi_{n-1}(\phi(\phi(s)))$$
$$= \phi_{n-2}(\phi(\phi(\phi(s))))$$
$$= \ldots\ldots\ldots\ldots\ldots\ldots$$
$$= \underbrace{\phi(\phi(\phi(\ldots(\phi(s))\ldots)))}_{n \text{ times}}$$
$$= \phi(\phi_n(s)), \text{ where } X_0 = 1 \qquad (4)$$

If $X_0 \neq 1$, then consider that $X_0 = k$. In that case, each of the k members will reproduce its offspring independently. These k members individually constitute their own branching processes. Thus, in this case, the generating function of the sum will be $[\phi_n(s)]^k$. Let us now try to find the expectations and variance.

Let us first assume that $X_0 = 1$, then considering the usual relationship and Eqn. (4), we get:

$$E(X_1) = \frac{d}{ds}\phi(s)\,|_{s=1} = \phi'(1) = \mu,$$ which is the mean number of offsprings of an individual.

$$E(X_2) = \frac{d}{ds}\phi_2(s)\,|_{s=1} = \frac{d}{ds}\phi(\phi(s))\,|_{s=1}$$
$$= \phi'(\phi(s))\phi'(s)\,|_{s=1}$$
$$= \{\phi'(1)\} = \mu^2$$

Proceeding in this manner, for $n \geq 1$,

$$E(X_{n+1}) = \frac{d}{ds}\phi_{n+1}(s)\,|_{s=1}$$
$$= \frac{d}{ds}\phi_n\,(\phi(s))\,|_{s=1}$$

$$= \phi_n' (\phi(s))\phi'(s)|_{s=1}$$
$$= \phi_n' (1)\phi'(1) = E(X_n)E(X_1) = \mu E(X_n)$$

Using induction, we get $E(X_{n+1}) = \mu^{n+1}$, $n = 1, 2, \ldots$. (5)

Note that, when $X_0 = k \neq 1$, then $E(X_{n+1} | X_0 = k) = k\mu^{n+1}$ (6)

Similar computations will yield, for variance, $n \geq 1$,

$$\begin{aligned} \text{Var}(X_{n+1}) &= E(X_{n+1}^2) - E(X_{n+1})^2 \\ &= E(X_{n+1}(X_{n+1}-1)) + E(X_{n+1}) - (EX_{n+1})^2 \\ &= \frac{d^2}{ds^2}\phi_{n+1}(s)|_{s=1} + \mu^{n+1} - (\mu^{n+1})^2 \\ &= \frac{d}{ds}[\phi_n'(\phi(s))\phi'(s)|_{s=1}] \\ &= \phi_n''(\phi(s))(\phi'(s))^2 + \phi_n'(\phi(s))\phi''(s) \\ &= \begin{cases} \sigma^2\mu^n(\mu^n - 1)/(\mu^2 - \mu), & \mu \neq 1 \\ n\sigma^2, & \mu = 1 \end{cases} \end{aligned} \tag{7}$$

where $\sigma^2 = \text{Var}\, X_1$.

Notice from Eqn. (5) that

$$E(X_{n+1}) \to \begin{cases} \infty; & \mu > 1 \\ 1; & \mu = 1 \\ 0; & \mu < 1 \end{cases}$$

In simple words, on the average, the generation size eventually **explodes** if $\mu > 1$, i.e., if the average number of offspring born to an individual is more than one, or, on average an individual is replaced by more than one offspring; similarly, the process dies out (becomes extinct) eventually if $\mu < 1$, i.e. if an individual is replaced, on the average, by less than one offspring. On the other hand, the average generation size remains fixed at 1, irrespective of the generation, if every individual is replaced by exactly one offspring, only on the average.

Here, it can be interpreted that the average of a branching process increases exponentially fast for super-critical, decreases exponentially fast for sub-critical and remains stable for critical.

Also, if $X_0 = k \neq 1$, then the variance is given below:

$$\begin{aligned} \text{var}(X_{n+1}) &= k\ \text{var}(X_{n+1} | X_0 = 1) \\ &= k\begin{cases} \sigma^2\mu^n(\mu^n - 1)/(\mu^2 - \mu), & \mu \neq 1 \\ n\sigma^2, & \mu = 1 \end{cases} \end{aligned}$$

Let us illustrate the concepts given above in the following examples.

Example 1: Consider a branching chain originating from a single element, where each individual element either replaces itself with probability p_1, or fails to replace itself (i.e., dies out without producing an offspring) with probability $p_0 = 1 - p_1$, $0 < p_0 < 1$. Obviously, $\mu = 0(p_0) + 1(p_1) = p_1 < 1$.

$E(X_n) = \mu^n \to 0$ as $n \to \infty$.

Here, $\phi(s) = E(s^x) = s^0(p_0) + s(p_1) = p_0 + p_1 s$.

Example 2: Consider a branching chain originated by a single element, where each individual element either replaces itself by k offspring with probability p_k, or fails to replace itself (i.e., dies out without producing an offspring) with probability $p_0 = 1 - p_k$ where $0 < p_0 < 1$.

Obviously, $\mu = 0(p_0) + k(p_k) = k\,p_k$ and $E(X_n) = \mu^n = (kp_k)^n \to 0, 1$ or ∞ as $n \to \infty$, depending on whether $k\,p_k < 1$, or $= 1$, or > 1. Here, $\phi(s) = E(s^x) = s^0(p_0) + s^k(p_k) = p_0 + p_k s^k$.

Example 3: Consider a branching chain where each individual in a generation either generates 2 offspring with probability p_2, or replaces itself with probability p_1, or fails to replace itself with probability $p_0, p_0 + p_1 + p_2 = 1$ where $0 < p_i < 1, i = 0, 1, 2$. Obviously, $\mu = 0(p_0) + 1(p_1) + 2(p_2) = p_1 + 2p_2$; also, $\phi(s) = E(s^x) = s^0(p_0) + s(p_1) + s^2(p_2) = p_0 + p_1 s + p_2 s^2$. Now, the limiting behaviour of $E(X_n)$ will depend on the numerical values of p_0, p_1 and p_2.

Example 4: Consider the probability distribution of the number of offspring in a branching chain originated by a single individual, where each individual element generates three offspring, following the Poisson law, with an average $\lambda > 0$, which is given by

$$p_k = (\lambda^k / k!)e^{-\lambda},\ k = 0, 1, 2, \ldots$$

thus,
$$\begin{aligned}\phi(s) = E(s^x) &= s^0(p_0) + s(p_1) + s^2(p_2) + \cdots \\ &= \{s^0(\lambda^0) + s(\lambda) + s^2(\lambda^2/2) + \cdots\}\ e^{-\lambda} \\ &= e^{s\lambda}e^{-\lambda} \\ &= \exp\{-\lambda(1-s)\}.\end{aligned}$$

and $\mu = \lambda. E(x_n) = \lambda^n \to 0$, or 1, or ∞ according as $\lambda < 1$, or $= 1$, or > 1.

Now, try the exercises that follow.

E1) Give two real life situations of the branching process.

E2) In a branching process $\{X_n : n \geq 1\}$, show that $E(X_{n+1}^2 \mid X_n = k) = k\sigma^2 + k^2\mu^2$. Hence find variance.

E3) Let $\{X_n\}$ be a branching process, where the probability distribution of the numbers of offspring be geometric, with $p_n = P[\xi = n] = qp^n$, $n = 0, 1, 2, \ldots$ and $q = 1 - p$. Then, find the probability generating function of $\{X_n\}$.

So far, we have discussed some mathematical concepts concerning the branching process. In a branching process there may arise a case when no further branch is reproduced. This is known as extinction. This may happen at any generation.

In the following section, we shall study the probability of extinction of a branching process in detail.

4.3 PROBABILITY OF EXTINCTION

By extinction we mean the event that the process $\{X_n\}$ will consist of all zeros but except for the first finitely n. It may also be noted in this context that in view of the structure of the process, once $X_n = 0$ for some n, say $n = k$, then X_n is zero for all subsequent n, i.e. $n = k+1, k+2, \ldots$ as there will be nobody left in generation k to reproduce generation $k+1, k+2, \ldots$. Let us denote the event $\{X_n = 0\}$ by $A_n, n \geq 1$. Since $X_n = 0$ implies $X_{n+1} = 0$, $A_n \subset A_{n+1}, n \geq 1$, so that $q_n = P(A_n) \uparrow$ as $n \uparrow$.

Recall that we had taken $\{p_n\}$ as the offspring distribution of an individual,

$$p_n = P(\xi = n), \; n = 0, 1, 2, \ldots$$

If $p_0 = 1$, then trivially, $P(X_1 = 0) = P(A_1) = 1$ so that the system is trivially extinct. On the other hand, if $p_0 = 0$, then every individual element reproduces at least one offspring so that X_n, the size of the n-th generation, which essentially comprise of the offspring of the elements of the previous generation, can never be zero with positive probability, so that the process can never be extinct. Thus, in what we discuss below, we take $0 < p_0 < 1$.

$$\begin{aligned}
\therefore \; P \text{ (extinction of the process)} &= P(X_n = 0 \text{ for some } n) \\
&= P\left(\bigcup_{n=1}^{\infty} A_n\right) \\
&= P\left(\lim_{n\to\infty} A_n\right), \text{ since } A_n \uparrow \\
&= \lim_{n\to\infty} P(A_n) \\
&= \lim_{n\to\infty} P(X_n = 0) \\
&= \lim_{n\to\infty} \phi_n(0) \text{ since } \phi_n(s) = Es^{X_n} = \sum_{j=0}^{\infty} s^j P(X_n = j) \qquad (8)
\end{aligned}$$

Obviously, $\phi(s) = Es^{X_1} \; 0 \leq s \leq 1$ is non-decreasing, and whenever $p_0 > 0$,

$$q_1 = P(X_1 = 0) = \phi(0) = P_0 > 0$$

$$q_1 = P(X_2 = 0) = \phi_2(0) = \phi(\phi(0)) = \phi(P_0) = \phi(q_1)$$

And, $\phi(q_1) > \phi(0) = q_1$, since $q_1 = p_0 > 0$.

Again,

$$\begin{aligned}
q_3 = \phi_3(0) = \phi(\phi_2(0)) &= \phi(q_2) \\
&> \phi(q_1), \text{since } q_2 > q_1 \text{ from above} \\
&> \phi(0) = q_1;
\end{aligned}$$

By induction,

$$q_{n+1} = \phi(q_n)$$

So that, from Eqn. (8), the extinction probability

$$\begin{aligned}
\Pi_0 &= \lim_{n\to\infty} \phi_n(0) \\
&= \lim_{n\to\infty} q_{n+1} \\
&= \lim_{n\to\infty} \phi(q_n) \\
&= \phi\left(\lim_{n\to\infty} q_n\right), \text{ by continuity of } \phi \\
&= \phi(\Pi_0).
\end{aligned}$$

Thus, the extinction probability Π_0 is the solution of the equation

$$\Pi_0 = \phi(\Pi_0) \qquad (9)$$

We may now summarize the above discussion as follows:

Theorem 1: Let $X_0, X_1, X_2, \ldots$ be a Galton-Watson process driven by the offspring distribution $\{p_0, p_1, \ldots\}$. Let $X_0 = 1$ and $\phi(s) = \sum_{j=0}^{\infty} s^j p_j$ be the probability generating function. Then, the extinction probability Π_0 of the process is zero, or one, depending on whether $p_0 = 0$ or 1; otherwise, Π_0 is a solution of the equation $s = \phi(s),\ 0 \le s \le 1$.

Notice that $s = 1$ is a root of the equation $s = \phi(s)$; but there can be other roots too! Then, which solution of $s = \phi(s),\ 0 \le s \le 1$ will yield Π_0? We will get further insight about Π_0 from the following theorem.

Theorem 2 (Fundamental Theorem): Suppose $p_0 > 0$. Then Π_0 is the smallest positive root of $s = \phi(s)$. If further $p_0 + p_1 < 1$, then $\Pi_0 = 1$ if and only if $\mu \le 1$.

Proof: We note that

$$s = \phi(s) = \sum_{j=0}^{\infty} s^j p_j \ge s^0 p_0 = q_1 .$$

Now, assume, $s \ge q_n$, then:

$$\begin{aligned} q_{n+1} &= P(X_{n+1} = 0) \\ &= \sum_{j=0}^{\infty} P(X_{n+1} = 0, X_1 = j) \\ &= \sum_{j=0}^{\infty} P(X_{n+1} = 0 \mid X_1 = j) P(X_1 = j) . \end{aligned} \tag{10}$$

Now, because of the structure of the process, under $X_1 = k$, the random vector $(X_2, X_3, \ldots, X_{n+1})$ is distributed as the sum of k independent vectors, exactly one originating from each of the k elements of generation 1. So that under $X_1 = k$, the distribution of $(X_2, X_3, \ldots, X_{n+1})$ is the same as the distribution of the sum of k i.i.d vectors each being distributed like $(X_1, X_2, \ldots, X_n)$. In particular

$P(X_{n+1} = 0 \mid X_1 = j)$

$= P$ (Size of the n-th generation of the j elements of generation-1 is zero)

$= P(\bigcap_{k=1}^{j}$ {Size of n-th generation of k^{th} element of the first generation is zero})

$= \prod_{k=1}^{j} P$ (Size of n-th generation of k^{th} element of the first generation is zero)

$= \prod_{k=1}^{j} P(X_n = 0)$

$= \{P(X_n = 0)\}^j = q_n^j .$

Thus, from Eqn.(10),

$$\begin{aligned} q_{n+1} &= \sum_{j=0}^{\infty} q_n^j p_j \\ &\le \sum_{j=0}^{\infty} s^j p_j = \phi(s) = s. \end{aligned}$$

Thus, by induction it follows that

$$s \geq q_n \text{ , } n = 1, 2, 3, \ldots$$
$$\therefore s \geq \lim_{n\to\infty} q_n = \lim_{n\to\infty} P(A_n) = \Pi_0$$

Thus, Π_0 is the smallest of the roots of $s = \phi(s)$. This proves the first part of the theorem. Now, let $p_0 + p_1 < 1$, i.e., the number of offspring of an element can be more than one with positive probability. Then,

$$\begin{aligned}
\phi''(s) &= \frac{d^2}{ds^2}\phi(s) \\
&= \frac{d^2}{ds^2}\sum_{j=0}^{\infty} s^j p_j \\
&= \sum_{j=0}^{\infty} j(j-1)\, s^{j-2}\, p_j \\
&> 0, \quad s \in (0, 1).
\end{aligned}$$

Therefore, ϕ is strictly convex in $(0, 1)$; also, $\phi'(s) = \sum_{j=0}^{\infty} js^{j-1}\, p_j$

> 0, since $p_0 < 1$,

so that $\phi(s)$ is strictly increasing in $s \in (0, 1)$. We now consider two cases.

Case I: $\mu \leq 1$

Then, ϕ being convex, $\phi'(s) < \phi'(1) = \mu \leq 1 = \frac{d}{ds} s$, so that

$$\frac{d}{ds}(\phi(s) - s) < 0.$$

This implies that $\psi(s) = \phi(s) - s$ is strictly decreasing in $s \in (0, 1)$. Thus, for $0 < s < 1$

$$\psi(s) > \psi(1)$$

i.e. $\quad \phi(s) - s > \phi(1) - 1 = 0$

i.e. $\quad \phi(s) > s, \ 0 < s < 1$

$\Rightarrow \quad \phi(s) = s$ does not have a solution in $(0, 1)$.

So that the only solution of $\phi(s) = s$ is $s = 1$, since $s = 0$ is not a solution as $\phi(0) = p_0 > 0$. Thus, we have been able to show that if $\mu \leq 1$, then $\Pi_0 = 1$.

Case II: $\mu > 1$

When $\phi'(1) = \sum_{j=0}^{\infty} j\, s^{j-1}\, p_j \,|_{s=1} = \mu > 1$, by the continuity of ϕ', $\exists\, s_0$, such that $\phi'(s) > 1$ for $s_0 < s < 1$, then:

expanding $\phi(s)$ about s_0, then at $s \neq 1$ we get

$$1 = \phi(1) = \phi(s_0) + (1 - s_0)\phi'(s^*), \ s_0 < s^* < 1$$

$$\therefore \quad \frac{1 - \phi(s_0)}{1 - s_0} = \phi'(s^*) > 1$$

i.e. $\phi(s_0) - s_0 < 0$.

Let us now try to visualize the graph of $\phi(s)$. We know the following already:

(i) $\quad \phi(s) - s$ is continuous,

(ii) $\quad \phi(0) = p_0 > 0, \ \phi(1) = 1, \ \phi(s_0) < s_0$

(iii) $\quad \phi(s)$ is convex

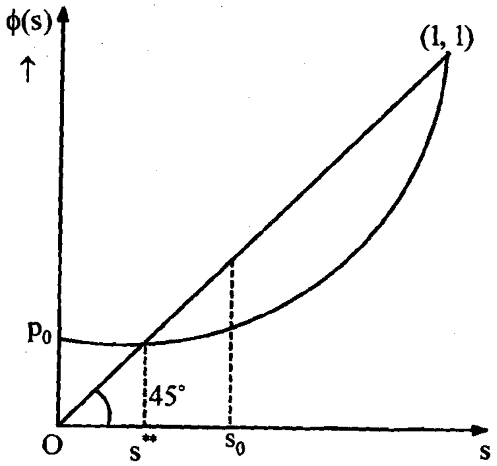


Fig.1

From (ii), then, $\exists\, s^{**} < s_0 \ni \phi(s^{**}) = s^{**}$. We will be through if we now show that $\phi(s)$ does not have any other root, obviously, smaller than s^{**}. If false, suppose $s^{***} < s^{**}$ is such a root so that $\phi(s^{***}) = s^{***}$, and $\psi(s) = \phi(s) - s = 0$ has three roots, viz, $s = s^{***}, s^{**}, 1$. Then, $\psi'(s) = \phi'(s) - 1 = 0$ has at least two roots in $(0, 1)$ and $\psi''(s) = \phi''(s) = 0$ has at least one root in $(0, 1)$, but $\phi''(s) = \sum_{j=0}^{\infty} j\,(j-1)s^{j-2}\,p_j > 0$ since $p_3 + p_4 + \cdots = 1 - (p_0 + p_1) > 0$ as $p_0 + p_1 < 1$. Thus, we confront a contradiction. Hence $\phi(s)$ has only one root in $(0, 1)$ .This completes the proof of the theorem.

Let us now illustrate it.

Example 5: Refer to Example 1. Here, $\phi(s) = E(s^x) = p_0 + p_1 s$, so that $s = 1$ is the only solution of the equation $\phi(s) = s$, as $p_0 + p_1 = 1$. As such, the probability Π_0 that the process will be extinct is 1.

Example 6: Refer to Example 2. Here, $\phi(s) = E(s^x) = s^0(p_0) + s^k(p_k) = p_0 + p_k s^k$; now, the equation $\phi(s) = s$ would imply that we need to solve $s - p_k s^k - p_0 = 0$, i.e., $s(1 - p_k s^{(k-1)}) = p_0$ for s. For $k > 1$, from Theorem 2, as $p_0 < 1$ (since $p_1 = 0$), the probability Π_0 of extinction of the process is 1, if and only, if $\mu = kp_k \le 1$.

Example 7: Refer to Example 4. Here, $\phi(s) = \exp\{-\lambda(1-s)\}$. Check that $p_0 + p_1 = e^{-\lambda} + \lambda e^{-\lambda} = (1+\lambda)e^{-\lambda} \le 1$ so that from Theorem 2, the probability Π_0 of extinction of the process is 1 provided $\lambda \le 1$; otherwise, $\Pi_0 < 1$.

Example 8: Let us consider the following offspring distribution:

$$p_j = b\,p^{j-1},\ \ j = 1, 2, \ldots$$

$$p_0 = 1 - \sum_{j=1}^{\infty} p_j = 1 - b/(1-p) = \frac{1-p-b}{1-p}$$

where $b, p > 0,\ \ p + b < 1$

During the 1920 census in the USA, this model was found to give a good fit with $b = 0.2126$ and $p = 0.5893$ in the context of American males. Now,

$$\phi(s)=\sum_{j=0}^{\infty} p_j\, s^j$$

$$=p_0+bs\sum_{j=1}^{\infty} p^{j-1}s^{j-1}$$

$$=p_0+\frac{bs}{1-ps}=\frac{1-p-b}{1-p}+\frac{bs}{1-ps}$$

$$\phi'(s)=\frac{b}{1-ps}+\frac{bs}{(1-ps)^2}.p$$

$$\therefore\ \ \mu=\phi'(s)\big|_{s=1}=\frac{b}{1-p}+\frac{bp}{(1-p)^2}$$

$$=\frac{b-bp+bp}{(1-p)^2}=\frac{b}{(1-p)^2}.$$

Also, $\phi(s)=s$ would mean $1-\dfrac{b}{1-p}+\dfrac{bs}{1-ps}=s$

i.e. $\quad b.\dfrac{-1+ps+s-ps}{(1-p)\,(1-ps)}=s-1$

i.e. $\quad \dfrac{b(s-1)}{(1-p)\,(1-ps)}=(s-1)$

i.e. $\quad \dfrac{b}{(1-p)\,(1-ps)}=1\ \text{ if }\ s\neq 1$

i.e. $\quad 1-ps=\dfrac{b}{1-p}$

i.e. $\quad s_0=\dfrac{1-p-b}{p(1-p)}$

Thus, 1 and s_0 are the two roots of $\phi(s)=s$. Note that

$$\mu \gtreqless 1$$

if $\quad b\gtreqless(1-p)^2$

which would mean

$$s_0=\frac{1-p-b}{p(1-p)}$$

$$\lesseqgtr\frac{1-p-(1-p)^2}{p(1-p)}$$

$$=1.$$

In summary,

$$\mu\gtreqless 1\Leftrightarrow s_0\lesseqgtr 1.$$

Thus, the extinction probability, according to Theorem 2, is $\Pi_0=1$ if $\mu\le 1$ and $\Pi_0=\dfrac{1-p-b}{p(1-p)}$ if $\mu>1$.

In this context, we may also be interested in determining the probabilities of extinction at the n-th generation. Here, algebra is interesting.

For any two points, μ and v,

$$\phi(s)-\phi(u)=b\left\{\frac{s}{1-ps}-\frac{u}{1-pu}\right\}=\frac{b(s-u)}{(1-ps)(1-pu)}$$

$$\therefore \quad \frac{\phi(s)-\phi(u)}{\phi(s)-\phi(v)} = \frac{s-u}{s-v} \cdot \frac{1-pv}{1-pu}.$$

Take $u = s_0$ and $v = 1$, so that for $\mu \neq 1$ (i.e. $s_0 \neq 1$)

$$\frac{1-p}{1-ps_0} = \left\{\frac{\phi(s)-s_0}{s-s_0}\right\}\left\{\frac{\phi(s)-1}{s-1}\right\}^{-1}$$

Taking the limit as $s \to 1$ (check for yourself)

$$\frac{1-p}{1-ps_0} = \frac{1}{\mu}$$

As such then, for all $s \neq 1$

$$\frac{\phi(s)-s_0}{\phi(s)-1} = \frac{1}{\mu} \cdot \frac{s-s_0}{s-1}.$$

Then, from Eqn. (3)

$$\frac{\phi_2(s)-s_0}{\phi_2(s)-1} = \frac{\phi(\phi(s))-s_0}{\phi(\phi(s))-1}$$

$$= \frac{1}{\mu} \cdot \frac{\phi(s)-s_0}{\phi(s)-1} = \frac{1}{\mu^2} \cdot \frac{s-s_0}{s-1}.$$

Iterating this way, for all $s \neq 1$

$$\frac{\phi_n(s)-s_0}{\phi_n(s)-1} = \frac{1}{\mu^n} \cdot \frac{s-s_0}{s-1}$$

which, when solved explicitly, yields, for $s \neq 1$

$$\phi_n(s) = 1 - \mu^n\left(\frac{1-s_0}{\mu^n - s_0}\right) + \frac{\mu^n\left(\dfrac{1-s_0}{\mu^n - s_0}\right)^2 s}{1-\left(\dfrac{\mu^n - 1}{\mu^n - s_0}\right)s}, \text{ if } \mu \neq 1. \tag{11}$$

If $\mu = 1$, then $s_0 = 1$ so that

$$\phi(s) = \frac{p-(2p-1)s}{1-ps}$$

as $b = (1-p)^2$.

Using Eqn. (3), it can be shown that for $n \geq 1$,

$$\phi_n(s) = \frac{np - \{(n+1)p - 1\}s}{1+(n-1)p - nps} \tag{12}$$

Thus, from Eqn. (11) and (12), the probability that the process will be extinct at generation n is

$$P(x_n = 0) = \phi_n(0)$$

$$= \begin{cases} 1-\mu^n\left(\dfrac{1-s_0}{\mu^n - s_0}\right), & \text{if } \mu \neq 1 \\ \dfrac{np}{1+(n-p)p}, & \text{if } \mu = 1 \end{cases}.$$

Now, try these exercises.

E4) Suppose in a branching process, the offspring distribution is as follows:
$p_k = ({}^N C_k) p^k q^{(N-k)}$, $q = 1-p$, $0 < p < 1$
$k = 0, 1, 2, \ldots$. Discuss the probability of extinction of this branching process.

E5) Following the notations introduced in this unit, for a branching process with $X_0 = 1$, define for $n = 1, 2, 3, \ldots$

$$Y_n = 1 + X_1 + X_2 + \cdots + X_n;$$

Thus, Y_n is the total number of individuals born in the family (progeny) up to and including the n-th generation. Let $\psi_n(s)$ denotes the probability generating function of Y_n. Show that

$$\psi_{n+1}(s) = s\phi(\psi_n(s)).$$

E6) Suppose in a branching process, the offspring distribution is as follows:

$$p_k = pq^k, \; q = 1-p, \; 0 < p < 1,$$

$k = 0, 1, 2, \ldots$. Discuss the probability of extinction in this branching process.

E7) Consider a branching chain where each individual element either replaces itself by 2 offspring with probability p, or fails to replace itself (i.e. dies out without producing an offspring) with probability $q = 1-p$, where $0 < p < 1$. We define $\psi_n(s)$ as in E5), and $\psi(s) = \lim_{n\to\infty} \psi_n(s)$. Show that if $p < 1/2$, then

$$\psi(s) = \frac{1-\sqrt{1-4pqs^2}}{2ps}.$$

What would be the expression for the limiting probability generating function $\psi(s)$ had the offspring distribution been geometric, as in E6).

We now end this unit by giving a summary of what we have covered in it.

4.4 SUMMARY

In this unit, we have discussed the following points.

1. The branching process is a Markov chain with a transition probability matrix $((P_{ij}))$.
2. The probability generation function of the offspring distribution $\{p_j\}$ is given by $\phi(s) = \sum_{j=0}^{\infty} s^j p_j; \; 0 \le s \le 1$.
3. The expectations and variance for the branching process are proved.
4. The extinction probability is given by $\Pi_0 = \phi(\Pi_0)$
5. The fundamental theorem for the probability of extinction is stated and proved.

4.5 SOLUTIONS/ANSWERS

E1) Any two real life situations constituting the branching process.

E3) The generating function

$$\phi(s) = \sum_j p_j s^j = \sum_{j=0}^{\infty} qp^j s^j = \frac{q}{1-ps}.$$

E4) The generating function

$$\phi(s) = \sum_j p_j s^j = \sum_{j=0}^{N} {}^N C_j \, p^j q^{N-j} s^j = (ps+q)^N$$

Now, the extinction probability has to be a solution of $s = (ps+q)^N$. Then, argue in terms of Theorems 1 and 2.

E5) Clearly, by definition,

$$\Psi_{n+1}(s)=\sum_{j=0}^{\infty} s^j P(Y_{n+1}=j)$$

$$=\sum_{j=1}^{\infty} s^j P(Y_n + X_{n+1}=j)$$

$$=\sum_{j=1}^{\infty} s^j \sum_{k=1}^{j} P(Y_n + X_{n+1}=j \mid Y_n=k)P(Y_n=k)$$

$$=\sum_{j=1}^{\infty} s^j \sum_{k=1}^{j} P(X_{n+1}=j-k \mid Y_n=k)P(Y_n=k)$$

$$=\sum_{k=1}^{\infty} s^k \sum_{j=k}^{\infty} P(\xi_{n1}+\xi_{n2}+\cdots+\xi_{nk}=j-k \mid Y_n=k)P(Y_n=k)s^{j-k}$$

where ξ_{nr} is the number of offspring born to the r-th member of generation n. The above line then equals to

$$=\sum_{k=1}^{\infty} s^k \sum_{j=k}^{\infty j} P(\xi_{n1}+\xi_{n2}+\cdots+\xi_{nk}=j-k)P(Y_n=k)s^{j-k}$$

$$=\sum_{k=1}^{\infty} s^k P(Y_n=k)\sum_{j=k}^{\infty} P(\xi_{n1}+\xi_{n2}+\cdots+\xi_{nk}=j-k)s^{j-k}$$

$$=\sum_{k=1}^{\infty} s^k P(Y_n=k)\sum_{r=0}^{\infty} P(\xi_{n1}+\xi_{n2}+\cdots+\xi_{nk}=r)s^{r}$$

$$=\sum_{k=0}^{\infty} s^k P(Y_n=k)\{\phi(s)\}^k \text{, since } \xi\text{'s are IID}$$

$$=\sum_{k=0}^{\infty} \{\phi(s)s\}^k P(Y_n=k)=\psi_n(s\phi(s)).$$

Alternately, let us look at

$$Y_{n+1}=1+X_1+X_2+\cdots+X_{n+1};$$

Thus, the total number of individuals born in the family till generation $n+1$, starting from generation 0 can be looked upon as the sum total of individuals born in n generations started by X_1 individuals of generation 1. Alternately, if $Y_1=k$, i.e., $X_1=k-1$, we should be able to write

$$Y_{n+1}=(1+X_{21}+\cdots+X_{n+1,1})+(1+X_{22}+\cdots+X_{n+1,2})+(1+X_{23}+\cdots+X_{n+1,3})$$
$$+\cdots+(1+X_{2\,k-1}+\cdots+X_{n+1,\,k-1})$$

where X_{ij} is the number of family members in (original) generation i of the branch started by the j-th individual (out of $k-1$ individuals) of the first generation. We also remember that the same stochastic conditions apply (reproduction distributions are the same for all individuals who behave independently of each other). Now,

$$\Psi_{n+1}(s)=\sum_{j=1}^{\infty} s^j P(Y_{n+1}=j)$$

$$=\sum_{j=1}^{\infty} s^j \sum_{k=1}^{j} P(Y_{n+1}=j \mid Y_1=k)\ P(Y_1=k)$$

$$=\sum_{j=1}^{\infty} s^j \sum_{k=1}^{j} P(Y_{n+1}=j \mid Y_1=k)P(Y_1=k)$$

$$=\sum_{j=1}^{\infty} s^j \sum_{k=1}^{j} P(1+X_{21}+\cdots+X_{n+1,1})+(1+X_{22}+\cdots+X_{n+1,2})$$
$$+(1+X_{23}+\cdots+X_{n+1,3})+$$
$$\cdots+(1+x_{2k-1}+\cdots+X_{n+1,k-1}=j)P(Y_1=k)$$

$$=\sum_{k=1}^{\infty}\sum_{j=k}^{\infty} s^j P((1+X_{21}+\cdots+X_{n+1,1})+(1+X_{22}+\cdots+X_{n+1,2})$$
$$+(1+X_{23}+\cdots+X_{n+1,3})+$$
$$\cdots+(1+X_{2k-1}+\cdots+X_{n+1,k-1})=j)P(Y_1=k)$$

$$=\sum_{k=1}^{\infty}\sum_{j=k}^{\infty} s^j P(Z_1+Z_2+\cdots+Z_{k-1}=j)\ P(Y_1=j)$$

where, for $r=1, 2, \ldots, k-1$, $Z_r=1+X_{2r}+\cdots+X_{n+1,r}$: here, Z_r is the number of family members born to the r-th member of the generation 1 till the $(n+1)$th generation counting from the first ever individual, and in n number of generations starting from the the r-th individual of the first generation. Notice that interpretation-wise, Z_r has a similar interpretation as that of Y_n, because Z_r is the number of family members caused in n generations by the r-th individual of the first generation. Thus, since the Z's are IID, being ascribed by different individuals of generation - 1, then the above is equal to (recalling that the pgf of sum of $(k-1)$ IID random variables is the $(k-1)$-th power of the pgf of any one of them)

$$=\sum_{k=1}^{\infty}\sum_{j=k}^{\infty} s^j \{P(Z_1+Z_2+\cdots+Z_{k-1}=j)\}P(Y_1=k)$$

$$=\sum_{k=1}^{\infty}\{\psi_n(s)\}^{k-1}P(X_1=k-1)$$

$$=\sum_{k=0}^{\infty}\{\psi_n(s)\}^{k}P(X_1=k)$$

$$\phi(s\psi_n(s)).$$

E6) $p_k=pq^k \quad q=1-p \quad 0<p<1 \quad p=0,1,2\ldots$

probability of extinction

$$\phi(s)=\sum_{j=0}^{\infty} p_j s^j$$

$$=\sum_{j=0}^{\infty} pq^j s^j$$

$$=\frac{p}{1-qs}$$

$$\phi'(s)=\frac{-p(-q)}{(1-qs)^2}=\frac{pq}{(1-qs)^2}$$

$$\mu=\phi'(s)|_{s=1}=\frac{pq}{(1-q)^2}$$

Also, $\phi(s) = s \Rightarrow \frac{p}{1-qs} = s$

i.e., $s_0 = p/q$.

Therefore 1 and s_0 are the two roots of $\phi(s) = s$.

It may also be noted that

$\mu \lesseqgtr 1$ if $q \lesseqgtr p$

i.e., $\mu \lesseqgtr 1$ if $s_0 \lesseqgtr p$.

Thus, the extinction probability is $\Pi_0 = 1$ if $\mu \le 1$, and $\Pi_0 = \frac{p}{1-p}$ if $\mu > 1$.

E7) Here, the offspring distribution is as follows:

$p_0 = q, p_1 = 0, p_{12} = p, p_r = 0, r \ge 3$

so that

$\varphi(s) = q + ps^2$.

Now, from E5), $\Psi_{n+1}(s) = s\varphi(\psi_n(s))$, and taking limit as $n \to \infty$ on both sides, we get

$\Psi(s) = s\varphi(\psi(s))$

and hence,

$\Psi(s) = s\{q + p\psi^2(s))$

i.e., $ps\Psi^2(s) - \psi(s)sq = 0$

so that $\Psi(s) = \frac{1-\sqrt{1-4pqs}^{\,2}}{2ps}$.

Whenever the offspring distribution is geometric, the pgf will be

$$\varphi(s) = \frac{p}{1-sq}$$

so that, arguing as before,

$$\psi(s) = \frac{ps}{1-q\Psi(s)}$$

and as such,

$q\Psi^2(s) - \psi(s) + ps = 0$.

Hence, $\Psi(s) = \frac{1-\sqrt{1-4pqs}}{2q}$.

---x---

UNIT 5 CONTINUOUS TIME MARKOV PROCESSES-I

Structure

5.1 INTRODUCTION

In the previous unit, we followed up on the size of the offspring population from one generation to another. Thus, it was as if we were dealing with the number of additions to the population through births over generations. In other words, X_n in the previous unit represented the number of births caused by the X_{n-1} elements of generation $(n-1)$, i.e,. X_n represented the number of descendents over a (generation) time period considered as one unit. It was as if we were counting time on a discrete scale. But there are phenomenon that experience continuous reproduction rendering the concept of a generation in their context inappropriate, and instead of looking at $\{X_n : n \geq 0\}$ one needs to look at $\{X_t : t \geq 0\}$ where X_t can be an alternative notation for $X(t)$ for convenience) is the number of 'births' or 'events' by time t or during the interval $[0, t]$ or alternately (equivalently) the 'population' at time t, depending on the context. This phenomenon can be seen in the following situations.

(i) Suppose, 'customers' arrive at a service point (like a ticket counter, or a doctor's chamber or barber shop, etc.) in a random fashion over time. We may denote $X(t)$ by the number of customers who arrived at the counter for service between time zero and t, i.e. by time t. It may be of interest, for instance, to study if the service facility is adequate or not in matching the arrivals.

(ii) Think of a machine of equipment which while working may fail from time to time. Suppose we have a very efficient maintenance system in place so that the moment it fails, it is repaired instantaneously so that the machine is once again functional. To work out a suitable maintenance strategy, it may be necessary to study the number of failures $X(t)$ that takes place between time zero and t, i.e., by time t.

We shall begin this unit by providing the necessary background on continuous time Markov Chains in Sec. 5.2. In Sec. 5.3, we give the mathematical formulation of continuous time Markov branching process in terms of the generating function and the mean function. We also provide a brief discussion on the extinction probabilities. In Sec. 5.4, we consider growth processes categorised as pure birth processes which are some variation of the branching process in continuous time. Here, we derive population size distribution at time t for Yule and Poisson processes besides some properties of the Poisson processes.

Objectives

After reading this unit, you should be able to:

- get the basic concepts of stochastic process in continuous time;
- use basic Kolmogorov forward/backward equations for obtaining population size distributions of special continuous time processes;
- obtain the generating function, the mean function of generation size at time t, and the extinction probability for a Markov branching process in continuous time;
- define a pure birth process and apply it to derive Poisson process and Yule process.

5.2 PRELIMINARIES

Before, we start continuous time Markov Process, let us discuss the prerequisites.

5.2.1 Continuous Time Markov Chain

Before we proceed to discuss 'Continuous Time Branching Processes" we need to recall certain concepts of Markov processes with discrete state space, which are also referred to as continuous time Markov chains. The objective is to prepare the ground for the introduction of Markovian continuous branching processes.

Consider a continuous time stochastic process $\{X(t): t \geq 0\}$ having a finite or countably infinite state space which we represent without any loss of generality by $S = \{0, 1, 2, \ldots\}$. Such a process is known as a **jump process**. If $X(t)$ **counts** the number of events (like number of births, number of failures of a machine, etc.) by time t, then $X(t)$ is called a **counting process**. Thus, for a process $\{X(t): t \geq 0\}$ to a counting process,

i) $X(t) \geq 0$ for all t

ii) $X(t)$ is integer-valued

iii) $X(s) \leq X(t)$ for $0 \leq s < t$, and

iv) $X(t) - X(s)$ for $0 \leq s < t$ would represent the number of events occurring in time interval $(s, t]$.

If for any process $\{X(t): t \geq 0\}$, $X(t_1)$, $X(t_2) - X(t_1)$, ..., $X(t_n) - X(t_{n-1})$ are independent random variables for all choices of $n \geq 2$ and $0 \leq t_1 < t_2 < \ldots < t_n$, then the stochastic process is said to have **independent increments**.

Define for $i, j \in S$, the transition probability function

$$P_{ij}(t) = P(X(t) = j \mid X(0) = i),\ t > 0,$$

$$P_{ij}(0) = \delta_{ij},$$

where δ_{ij} is the Kronecker's delta, i.e.

$$\delta_{ij} = \begin{cases} 1;\ i = j \\ 0;\ i \neq j \end{cases}$$

with $\sum_{j \in S} P_{ij}(t) = 1$. The matrix $(P_{ij}(t))$ is called the transition probability function matrix associated with the stochastic process $X(t), t \geq 0$. For a continuous time

Markov chain, we require

$$P(X(t)=j \mid X(t_1)=i_1,\ X(t_2)=i_2,\ldots,\ X(t_n)=i_n,\ X(s)=i)$$
$$= P(X(t)=j \mid X(s)=i) \qquad (1)$$

for every $n \geq 1$, $0 \leq t_1 < t_2 < \ldots < t_n < s < t$ and $i_1, i_2, \ldots, i_n, i, j \in S$. This is known as the **Markov property**. We shall deal with the case where the transition probability function is stationary with respect to the time, or time homogeneous, i.e., it depends only on the length of the time interval, but not on its location. Thus, in this case the right hand side of Eqn. (1) is essentially $P_{ij}(t-s)$. The Poisson process, or the branching process, that we shall discuss shortly, are examples of such a process.

Consider a Markovian stochastic process $\{X(t): t \geq 0\}$ with stationary transition probability function $(P_{ij}(t))$. Then, for $0 \leq t_1 < t_2 < \ldots < t_n$,

$$P(X(t_1)=i_1,\ X(t_2)=i_2,\ \ldots,\ X(t_n)=i_n)$$
$$= P(X(t_n)=i_n \mid X(t_{n-1})=i_{n-1})P(X(t_{n-1})=i_{n-1},\ X(t_{n-2})=i_{n-2},\ \ldots,\ X(t_1)=i_1)$$
$$= Pi_{n-1}i_n(t_n - t_{n-1})Pi_{n-2}i_{n-1}(t_{n-1}-t_{n-2})\ldots Pi_1i_2(t_2-t_1)P(X(t_1)=i_1).$$

thus,

$$P(X(t)=j,\ X(t+h)=k \mid X(0)=i) = P(X(t)=j,\ X(t+h)=k,\ X(0)=i)/P(X(0)=i)$$
$$= P(X(t+h)=k \mid X(t)=j)\ P(X(t)=j \mid X(0)=i)$$
$$= P_{jk}(h)\ P_{ij}(t),$$

thus, for any $t, h \geq 0$, we have

$$P_{ij}(t+h) = P(X(t+h)=j \mid X(0)=i)$$
$$= \sum_{k \in S} P(X(t+h)=j, X(t)=k \mid X(0)=i)$$
$$= \sum_{k \in S} P_{ik}(t)\ P_{kj}(h). \qquad (2)$$

This is the **Chapman-Kolmogorov** equation. In matrix notation, this equation can be represented as

$$\mathbf{P}(t+h) = \mathbf{P}(t)\mathbf{P}(h) \qquad (3)$$

where $\mathbf{P}(t) = (P_{ij}(t))$ and $P(0) = (S_{ij}) = 1$ by definition. Notice that any t can be expressed in relation to another number $t_0 < t$ as $t = r\ t_0 + s$, where r is a positive integer and $0 \leq s < t_0$, so that

$$\mathbf{P}(t) = \mathbf{P}(r\ t_0 + s)$$
$$= \mathbf{P}(t_0)\ \mathbf{P}((r-1)t_0 + s)$$
$$= \mathbf{P}^2(t_0)\ \mathbf{P}((r-2)t_0 + s)$$
$$= \ldots$$
$$= \mathbf{P}^r(t_0)\ \mathbf{P}(s). \qquad (4)$$

This result has a deep significance; it shows that once $\mathbf{P}(t)$ is specified for $0 < t \leq t_0$, however small t_0 may be, we can use the Chapman-Kolmogorov equation to know $\mathbf{P}(t)$ for all $t > 0$.

Now, let us discuss Kolmogorov forward and backward equations in detail.

5.2.2 Kolmogorov Forward and Backward Equations

Let us assume that, as is logical, for all $i, j \in S$

$$\lim_{(t \to 0)} P_{ij}(t) = \delta_{ij}$$

In matrix notation,

$$\lim_{(t \to 0)} \mathbf{P}(t) = \mathbf{I}$$

Note that we have also defined

$$\mathbf{P}(0) = \mathbf{I}.$$

Also, let us assume that $P_{ij}(t)$ are continuous for $t > 0$. Then, under the above assumptions, it can be shown that $\mathbf{P}(t)$ is differentiable at all $t \geq 0$ (of course, the derivative at $t = 0$ is the right hand derivative). Define for all $i, j \in S$

$$Q_{ij} = P_{ij}(0) = \lim_{(t \to 0)} \frac{\{P_{ij}(t) - P_{ij}(0)\}}{t}.$$

These are well-defined. The quantities Q_{ij} are called the **infinitesimal transition rates** and $\mathbf{Q} = (Q_{ij})$ is called the **infinitesimal generator or infinitesimal rate matrix.**

Whenever the state space S is finite, then, differentiating both sides of Eqn. (2) with respect to t at $t = 0$, we get

$$P'_{ij}(h) = \sum_{k \in S} P'_{ik}(0)\, P_{kj}(h)$$

$$= \sum_{k \in S} Q_{ik}\; P_{kj}(h),$$

so that for all $t \geq 0$.,

$$\mathbf{P}'(t) = \mathbf{Q}\,\mathbf{P}(t) \tag{5}$$

This is called **Kolmogorov's Backward Equation.** Again, differentiating Eqn. (2) with respect to h and setting $h = 0$, we get

$$P'_{ij}(t) = \sum_{k \in S} P_{ik}(t)\, P'_{kj}$$

$$= \sum_{k \in S} P_{ik}(t)\, Q_{kj}$$

so that for all $t \geq 0$,

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}, \tag{6}$$

which is **Kolmogorov's Forward Equation.**

It is interesting to note here that whenever S is finite and $\mathbf{P}(0) = \mathbf{I}$, the backward Eqn.(5) with the initial condition $\mathbf{P}(0) = 1$, has a unique solution $\mathbf{P}(t) = \exp(t\mathbf{Q})$, where we define

$$\text{ext}(t\mathbf{Q}) = \mathbf{I} + \sum_{n=1}^{\infty} t^n \mathbf{Q}^n / n!.$$

as a format matrix series.

Now we are ready to discuss continuous time Markov branching process in this section.

5.3 CONTINUOUS TIME MARKOV BRANCHING PROCESSES

We shall start out by a explaining a popular notation regarding a real-valued function $f(.)$ defined on the real line. We shall write "$f(x) = 0(x)$ as $x \to 0$", read as "the function $f(x)$ is lower order as x goes to zero" to mean that $f(x)$ is such that $f(x)/x \to 0$ as $x \to 0$. For instance, $f(x) = x^{1.2} = 0(x)$ as $x \to 0$ while $f(x) = x \neq 0(h)$ as $x \to 0$. It is not hard to see that:

- If $f(x) = 0(x)$ as $x \to 0$ and $g(x) = 0(x)$ as $x \to 0$, then $f(x) + g(x) = 0(x)$ as $x \to 0$.
- If $f(x) = 0(x)$ as $x \to 0$, then for any scalar γ, $\gamma f(x) = 0(x)$ as $x \to 0$.

We will have to make use of these two trivial results in our succeeding computations.

Let us formulate the Markov branching process in continuous time mathematically.

Consider a continuous time Markovian process $\{X(t) : t \geq 0\}$, where $X(t)$ denotes the number of elements at time t given $X(0) = 1$. Each element, after surviving for a random amount of time, splits upon producing a random number of elements (offspring) of identical type; we assume the process to be time homogeneous in the sense that an element splits upon producing j elements over a time duration $(t, t+h], h > 0$ with probability

$$\delta_{1j} + \alpha_j h + 0(h), \quad j = 0, 1, 2, \ldots \tag{7}$$

as $h \to 0$, where δ_{1j} is the Kronecker delta, i.e., $\delta_{1j} = 1$, or 0 accordingly as $j = 1$ or $j \neq 1$. It is assumed that $\alpha_j \geq 0$ for $j \neq 1$, and $\sum_{j=0}^{\infty} \alpha_j = 0$. We further assume that individual elements act independently of each other in accordance with probabilities given in Eqn. (7). We observe that for $j \neq 1$

$$\begin{aligned} P_{n\,n+j-1}(h) &= P\{X(t+h) = n+j-1 \mid X(t) = n\} \\ &= n\{\alpha_j h + 0(h)\}\{1 + \alpha_1 h + 0(h)\}^{n-1} + 0(h) \\ &= n\,\alpha_j h + 0(h), \end{aligned} \tag{8}$$

while

$$\begin{aligned} P_{nn}(h) &= P\{X(t+h) = n \mid X(t) = n\} \\ &= \{1 + \alpha_1 h + 0(h)\}^n + 0(h) \\ &= 1 + n\alpha_1 h + 0(h). \end{aligned} \tag{9}$$

Let us pause for a while to ponder how these expressions could be arrived at from Eqn. (7). At time point t, $X(t) = n$, i.e., the number of elements is n at time t, let us explore the ways the event that the same can be $n + j - 1$ after h units of time. It can happen in many ways, for instance, suppose one of the elements splits into j elements and each of the remaining $(n-1)$ elements is replaced by exactly one offspring so that at time $t + h$ the number of elements would be $j + n - 1 = n + j - 1$. Since each element behaves independently of the other, we are in a binomial situation and the probability of this particular possibility is exactly equal to the first expression in Eqn. (8). But the same event can happen also if, say, one of them is replaced by $j - 1$ elements, another by 2 elements, and that each of the remaining $(n-2)$ elements is replaced by exactly one offspring, so that at time $t + h$, $X(t+h) = (j-1) + 2 + (n-2) = n + j - 1$. But the probability of the later will be, by independence again, is given by $n(n-1)\{\alpha_{j-1} h + 0(h)\}\{\alpha_2 h + 0(h)\}\{1 + \alpha_1 h + 0(h)\}^{n-2}$ the whole of which is $0(h)$ as $h \to 0$. In the same way, all other possibilities can be shown to have probabilities which together will be of $0(h)$ as $h \to 0$. This explains the expression given in Eqn. (8).

Let us now examine Eqn. (9). One of the ways in which the number of elements remain the same at n both at time points t and $t + h$ is when each of the n elements at time point t is replaced by an element produced by it and the relevant probability is $\{1+\alpha_1 h+0(h)\}^n = 1+n\alpha_1 h+0(h)$. Any other possibility of achieving the event of the number of elements to be pegged at n at both time points t and $t+h$ will be of $0(h)$ as $h \to 0$.

Let

$$\phi(t;\, s) = \sum_{j=0}^{\infty} s^j\, P_{1j}(t) = E\, s^{x(t)} \tag{10}$$

the probability generating function of $P_{1j}(t) = P\{X(t)=j \mid X(0)=1\}$, $j = 0, 1, 2, \ldots$; trivially, $\phi(0;\, s) = s$. Also, because of independence amongst the elements, using arguments as in Unit 4, for $i \geq 0$, the probability generating function of $P_{ij}(t)$ can be seen to be

$$\sum_{j=0}^{\infty} s^j\, P_{ij}(t) = \{\phi(t;\, s)\}^i\, . \tag{11}$$

At this stage, let us recall the Chapman-Kolmogorov Eqn. (2) as follows

$$P_{ij}(t+h) = \sum\nolimits_{k\in S} P_{ik}(t)\, P_{kj}(h)\, .$$

Now, from Eqn. (11)

$$\begin{aligned}\{\phi(t+h;\, s)\}^i &= \sum_{j=0}^{\infty} s^j\, P_{ij}(t+h)\\ &= \sum_{j=0}^{\infty} s^j \sum_{k=0}^{\infty} P_{ik}(t)\, P_{kj}(h)\\ &= \sum_{k=0}^{\infty} P_{ik}(t) \sum_{j=0}^{\infty} s^j\, P_{kj}(h)\\ &= \sum_{k=0}^{\infty} P_{ik}(t)\{\phi(h;\, s)\}^k\\ &= \{\phi(t;\, \phi(h;\, s))\}^i\, ,\end{aligned} \tag{12}$$

and with $i = 1$,

$$\phi(t+h;\, s) = \phi(t;\, \phi(h;\, s))\, , \tag{13}$$

a result equivalent to Eqn. (4) in Unit 4 on the discrete case.

Let us now specialize these results in the context of our present model as specified by Eqn. (7). Let us denote it by $u(s) = \sum_{j=0}^{\infty} \alpha_j\, s^j$. Then

$$\begin{aligned}\phi(h;\, s) = \sum_{j=0}^{\infty} s^j\, P_{1j}(h) &= \sum_{j=0}^{\infty} s^j\, (\delta_{1j} + \alpha_j\, h + 0(h))\\ &= s + h\, u(s) + 0(h)\, ,\end{aligned} \tag{14}$$

so that from Eqn. (13),

$$\phi(t+h;\, s) = \phi(t;\, s + h\, u(s) + 0(h))\, .$$

Now, we shall use the Taylor's expansion of the bivariate function ϕ as below

$$\begin{aligned}\phi(t+h;\, s) &= \phi(t;\, s + h\, u(s) + 0(h))\\ &= \phi(t;\, s) + h\, u(s)\, \{d\phi(t;\, s)/ds\} + 0(h),\end{aligned}$$

and, thus,

$$\{\phi(t+h;\, s) - \phi(t;\, s)\}/h = \{d\phi(t;\, s)/ds\; u(s) + 0(h)/h\, .$$

Taking the limit of both sides as $\lambda \to 0$, we have,

$$\partial\phi(t;\, s) = \{\partial\phi(t;\, s)\}\ u(s)\,. \tag{15}$$

This partial differential equation can be solved for $\phi(t;\, s)$ subject to the initial condition $\phi(0;\, s) = s$.

Again, from Eqn. (13) and Eqn. (14), we get

$$\begin{aligned}\phi(t+h;\, s) &= \phi(t;\, \phi(h;\, s)) \\ &= \phi(h;\, s) + t\ u(\phi(h;\, s)) + 0(t)\end{aligned}$$

Now interchanging t and λ, the same equation can be written as

$$\phi(t+h;\, s) = \phi(t;\, s) + h\ u(\phi(t;\, s)) + 0(h)\,.$$

Hence,

$$\{\phi(t+h;\, s) - \phi(t;\, s)\}/h = u(\phi(t;\, s)) + 0(h)/h$$

so that by taking the limit as t tends to zero, we have another equation, this time an ordinary differential equation

$$\frac{\partial}{\partial t}\phi(t;\, s) = u(\phi(t;\, s)) \tag{16}$$

which again can be solved under the initial condition $\phi(0,\, s) = s$. The Eqn. (15) and Eqn. (16) may be referred to as forward and backward equations, respectively.

We wish to evaluate the mean function, $m(t) = EX(t)$, under the above set up. As we already know, a convenient way to compute this would be to differentiate the probability generating function $\phi(t;\, s)$ with respect to s, and to evaluate the same at $s = 0$. In this context, we first differentiate both sides of Eqn. (15) with respect to s, and interchange the order of differentiation, to get

$$\frac{\partial}{\partial s}\left\{\left(\frac{\partial}{\partial t}\right)\phi(t;s)\right\} = \frac{\partial}{\partial s}\left\{u(s)\left(\frac{\partial}{\partial s}\right)\phi(t;s)\right\}$$

i.e. $$\frac{\partial}{\partial t}\left\{\left(\frac{\partial}{\partial s}\right)\phi(t;s)\right\} = u(s)\frac{\partial^2}{\partial s^2}\{\phi(t;s)\} + \left(\frac{\partial}{\partial s}\right)\{\phi(t;s)\}\,u'(s).$$

For $s = 1, u(1) = 0$, we have

$$\frac{\partial}{\partial t}\ m(t) = u'(1)\ m(t)\,.$$

Under the initial condition $X(0) = 1$, this can be easily solved for m(t) to yield

$$m(t) = e^{u'(1)\,t}\,. \tag{17}$$

Now, we shall discuss the probability of extinction of this process.

The process cannot be extinct unless $\alpha_0 > 0$. And, as such it will make sense to consider only the case $\alpha_0 > 0$. Also, we take $X(0) = 1$ as this has been argued earlier in Eqn. (11)

$$\sum_{j=0}^{\infty} s^j\, P_{ij}(t) = \{\phi(t;\, s)\}^i = \left\{\sum_{j=0}^{\infty} s^j\, P_{1j}(t)\right\}^i$$

so that, $P_{i0}(t) = \{P_{10}(t)\}^i$. Here $P_{i0}(t)$ is the probability of a population of size i dying out by time t. Notice that Eqn. (13) and the fact that $\phi(t;s)$ is non-decreasing for fixed t show that

$$P_{i0}(t+h) = \{P_{10}(t+h)\}^i = \{\phi(t+h;\, 0)\}^i = \{\phi(t;\, \phi(h;\, 0))\}^i \ge \{\phi(t;\, 0)\}^i = P_{i0}(t)\,,$$

so that $P_{i0}(t)$ is non-decreasing in t, which is intuitively clear as well. Given this background, the "extinction probability", the probability that a family caused by an individual will eventually die out, can be defined as $\pi_0 = \lim\limits_{t\to\infty} P_{10}(t)$.

We shall close this section after quoting the following theorem without proof which is a bit involved, but not hard:

Theorem 1: Suppose $\alpha_0 > 0$. Then, π_0 the smallest non-negative root of the equation $u(s) = \sum_{j=0}^{\infty} \alpha_j s^j = 0$. Further $\pi_0 = 1$ if and only if $u'(1) \le 0$.

Now, let us apply this theorem in the following example:

Example 1: Consider a continuous time Markov branching process with $\alpha_0 = a, \alpha_1 = -(a+b), \alpha_2 = b$, and $\alpha_j = 0$ for $j \ge 3$.

Then

$$\phi(s) = \sum_{j=0}^{\infty} \alpha_j s^j$$

$$\text{or } \phi(s) = bs^2 - (a+b)s + a$$

$$= b(s - a/b)(s-1)$$

Here, it is clear that $\pi_0 = 1$ iij $b \le a$

or $\pi_0 = 1$ iij $\phi'(1) \le 0$.

5.4 PURE BIRTH PROCESSES

We will now consider a variation of the Markov branching process discussed in the previous section. In that process, we postulate as follows:

$$P\{X(t+h) = k+1 \mid X(t) = k\} = k\alpha h + g_1(h)$$

$$P\{X(t+h) = k \mid X(t) = k\} = 1 - k\alpha h + g_2(h)$$

and for $j \ne 0,1$

$$P\{X(t+h) = n+j \mid X(t) = n\} = g_j(h)$$

for some $\alpha > 0$, where $g_j(h)$ are $o(h)$ as $h \to 0$, $j = -1, 0, 1, 2, \ldots$. Note the characteristics of this process. Here, it is clear that the chances of a 'birth' in the time interval $[t, t+h]$ depends on the size of the population as well as on the length of the time interval and further, those of more than one 'birth' in the time interval $[t, t+h]$ is negligible. No death is allowed.

To generalize further, suppose $\{X(t) : t \ge 0\}$ is a Markov process, and for series of positive numbers λ_k, the following postulates are satisfied as $h \to 0$.

1. $P_{kk+1}(h) = P\{X(t+h) = k+1 \mid X(t) = k\} = \lambda_k h + g_1(h)$
2. $P_{kk}(h) = P\{X(t+h) = k \mid X(t) = k\} = 1 - \lambda_k h + g_2(h)$
3. $P\{X(t+h) = k+j \mid X(t) = k\} = g_j(h)$, for all $j = -1, 2, 3, \ldots$ where $g_j(h) = 0(h)$ as $h \to 0$.

   In addition, sometimes, to be able to handle the consequent algebra and also to be in tune with the reality of the process being modeled, when $\lambda_0 > 0$ we may postulate the following:
4. $X(0) = 0$, for convenience.

This is required to be used as the initial condition at times and also postulate 4 may be changed if felt appropriate in a given situation. Notice that in view of the last postulate, $X(t)$ can be interpreted as the number of births in $[0, t]$ and not the size of the population at time t. Also, postulates 1 and 2 together imply that the chance of more than one birth over a period of duration h is of lower order h. Such processes can be referred to as pure birth processes. Let $P_n(t) = P(X(t) = n)$. Then, for $n \ge 1$,

$$P_n(t+h) = P(X(t+h)=n) = \sum_{j=0}^{\infty} P(X(t+h)=n,\ X(t)=j)$$

$$= P(X(t+h)=n,\ X(t)=n-1) + P(X(t+h)=n,\ X(t)=n) + 0(h)$$

$$= P(X(t+h)=n \mid X(t)=n-1)\,P(X(t)=n-1)$$

$$+ P(X(t=h)=n \mid X(t)=n)P(X(t)=n + 0(h)$$

$$= \{\lambda_{n-1}h + g_1(h)\}\,P_{n-1}(t) + \{1-\lambda_n h + g_2(h)\}\,P_n(t) + 0(h)$$

so that

$$\{P_n(t+h) - P_n(t)\}/h = \lambda_{n-1}P_{n-1}(t) - \lambda_n P_n(t) + \frac{0(h)}{\lambda}.$$

Taking the limit as $h \to 0$, we get

$$P_n'(t) = \lambda_{n-1}\,P_{n-1}(t) - \lambda_n\,P_n(t) \tag{18}$$

On the other hand,

$$P_0(t+h) = P(X(t+h)=0) = P(X(t+h)=0,\ X(t)=0)$$

$$= P(X(t+h)=0 \mid X(t)=0)\ P(X(t)=0)$$

$$= \{1-\lambda_0 h + g_2(h)\}\ P_0(t)$$

so that, as before, we now get

$$P_0'(t) = -\lambda_0\,P_0(t) \tag{19}$$

Solving Eqn. (19) with the initial condition $X(0)=0$, i.e., $P_0(0)=1$,

$$P_0(t) = \exp(-\lambda_0 t) \tag{20}$$

Notice that since $X(0)=0$ as per our postulate 4, the event $\{X(t)=0\} = \{T_0 > t\}$, where T_0 is the random variable denoting the time to the first birth; according to Eqn. (20) then, $P_0(t) = P(X(t)=0) = P\{T_0 > t\}$, so that the time to the first event is exponentially distributed with the mean $(1/\lambda_0)$.

Now, from Eqn. (17), solving successively, for $n = 1, 2, \ldots$

$$P_n(t) = \lambda_{n-1}\exp(-\lambda_n t)\int_0^t \exp(\lambda_n u)\ P_{n-1}(u)\,du\,. \tag{21}$$

Example 2 (Yule Process): Consider a pure birth process that has exactly one element originally (originator). It and its successive offspring each has a probability $\alpha h + 0(h)$ of reproducing a similar element over a period of infinitesimal length $h(\alpha > 0)$; as usual, we assume that the elements reproduce independently of each other. Remember that we are not bringing in the concept of "death" and as such the value of $X(t)$ can only go up through births at random points of time. Also, note that given that the probability of birth over a duration of length h is $\alpha h + 0(h)$, the probability of no birth is $(1-\alpha h + 0(h))$ as more than one birth is not permitted. Suppose that the population has achieved size k at time t, i.e., $X(t)=k$, and after h units of time its size has to become, say, $k+1$, there being no 'deaths', the only way it can happen would be when exactly one out of k elements present must have reproduced an offspring (the probability of which for each of them is $\alpha h + 0(h)$) while the others did not (the probability of which for each of them is $1-\alpha h + 0(h)$). Thus, because of independence of the behaviour of the k elements, we are in a binomial situation, and obviously,

$$P(X(t+h)=k+1 \mid X(t)=k) = ({}^kC_1)\{\alpha h + 0(h)\}^1\{1-\alpha h + 0(h)\}^{k-1}$$

$$= k\lambda h + 0(h).$$

Arguing in a similar fashion, for $j > 1$,

$$P(X(t+h)=k+j \mid X(t)=k) = ({}^kC_j)\ \{\alpha h + 0(h)\}^j (1-\alpha h + 0(h))^{k-j}$$
$$= 0(h).$$

Thus, comparing this with the model discussed just before Example 2, we find that $\lambda_k = k\lambda$. The birth process with parameters $\lambda_k = k\lambda$ is called a **Yule Process**. The solutions of Eqns. (19) and (20) through routine computations with the initial conditions, $X(0)=1$ (i.e. $P(X(0)=j)=\delta_{1j}$, $j = 0, 1, 2, \ldots$) yields

$$P_k(t) = e^{-\alpha t}(1-e^{-\alpha t})^{k-1},\ k = 1, 2, \ldots \qquad (21)$$

We note that for each $t > 0$, $X(t)$ has a geometric distribution with success probability $e^{-\alpha t}$. Therefore, $EX(t) = (1/e^{-\alpha t}) = e^{\alpha t}$, i.e. on the average, the process grows exponentially.

Example 3 (Poisson Process): The postulates of a pure birth process may be recounted as follows:

In the pure birth process we allowed the probability of a birth over a small interval of time to depend on the number of births k. But, suppose that, in a special situation, the λ_k's are identical and equal to a constant λ. For instance, $X(t)$ can be the number of occurrences of an event of a certain kind taking place in the time interval $[0, t]$; e.g. $X(t)$ can be the number of accidents taking place in a city over time, or number of telephone calls passing through a PABX board, or number of times a particular component of a machine is replaced following failures over time, or even the number of customers arriving at a store over a time, the number of cars that pass a particular bus-stop by time t, etc. In such cases, obviously, no event lies in the root of the occurrence of a previous one (i.e. unlike fathering the "next generation") so that the number of events that have taken place until the recent time will not have any apparent influence whatsoever on the occurrence of a subsequent one; at the same time it is logical to argue that the chance of occurrence of event in a time interval of small length is proportional to its length. We will later approach the Poisson process from this aspect as well.

We define a Poisson process as a birth process satisfying the following postulates:

1. $P_{k\ k+1}(h) = P\{X(t+h)=k+1 \mid X(t)=k\} = \lambda h + g_1(h)$
2. $P_{kk}(h) = P\{x(t+h)=k \mid X(t)=k\} = 1-\lambda h + g_2(h)$
3. $P\{X(t+h)=k+j \mid X(t)=k\} = g_j(\lambda)$, for all $j \neq 0,1$
4. $X(0)=0$

where $g_i(h) = 0(h)$ as $h \to 0$.

Then, from Eqn. (17) and Eqn. (18), we have

$$P_0'(t) = -\lambda P_0(t)$$
$$P_n'(t) = \lambda P_{n-1}(t) - \lambda P_n(t),$$

the initial conditions being, because of postulate 4 given above, $P_0(0)=1, P_n(0)=0$ for $n \geq 1$. Then from Eqn. (19), we will have

$$P_0(t) = \exp(-\lambda t) \qquad (22)$$

and for $n \geq 1$

$$P_n(t) = \lambda \exp(-\lambda t)\int_0^t \exp(\lambda u)\ P_{n-1}(u)\ du \qquad (23)$$

which can be recursively solved to yield

$$P_n(t) = P(X(t) = n) = \frac{(\lambda t)^n}{n!}\exp(-\lambda t),$$

So that for $t > 0$, $X(t)$, the number of event occurrences of events that take place up to time t, have a Poisson distribution with mean λt. Is there an alternate set of postulates to identify a Poisson process? In fact, we shall now show that a counting process, $\{X(t): t \geq 0\}$ ($X(t)$, that often denotes the number of births in the time interval $[0, t]$) is a Poisson process with rate $\lambda > 0$, if

1. $X(0) = 0$
2. $\{X(t); t > 0\}$ has independent increments
3. for any $t > 0$, $X(t+h) - X(t)$, which is the number of events (births) in the interval $(t, t+h]$ of length h follows a Poisson distribution with mean λh.

Notice that because of 2 and 3,

$$\begin{aligned} P_{kk+1}(h) = P(X(t+h) = k+1 \mid X(t) = k\} &= P\{X(t+h) - X(t) = 1 \mid X(t) = k\} \\ &= P\{X(t+h) - X(t) = 1\} \\ &= \lambda h e^{-\lambda h} \\ &= \lambda h\{1 - \lambda h + (\lambda h)^2/2 - \ldots\} \\ &= \lambda h + g_1(h) \\ &= \lambda h + 0(h) \end{aligned}$$

since

$|g_1(h)| \leq (\lambda h)^2\{1 + (\lambda h)/2 + (\lambda h)^2/3! + (\lambda h)^3/4! + \cdots\} \leq (\lambda h)^2\{1 + (\lambda h) + (\lambda h)^2/2! + (\lambda h)^3/3! + \cdots\} = (\lambda h)^2 e^{\lambda h}$ so that $|g_1(h)/h| \leq (\lambda^2 h)e^{\lambda h}$; thus $g_1(h) = 0(h)$ as $h \to 0$. This shows that the postulate 1 of the Poisson process is satisfied, i.e., $P_{kk+1}(h) = P\{X(t = h) = k+1 \mid X(t) = k\} = \lambda h + 0(h)$.

Next, let us note that

$$P\{X(t+h) - X(t) \geq 2\} = \sum_{n=2}^{\infty} e^{-\lambda h}(\lambda h)^n / n! \leq e^{-\lambda h}(\lambda h)^2 \sum_{m=0}^{\infty} (\lambda h)^m / m! \leq (\lambda h)^2$$

which is $0(h)$ as $h \to 0$ justifying postulate 3. Finally

$$\begin{aligned} P_{kk}(h) &= P\{X(t+h) - X(t) = 0\} = P(X(t+h) - X(t) = 0\} \\ &= 1 - P\{X(t+h) - X(t) = 1\} - P(X(t+h) - X(t) \geq 2\} = 1 \\ &\quad -\{(\lambda h) + 0(h)\} - 0(h) = 1 - (\lambda h) + 0(h), \end{aligned}$$

so that postulate 2 is also satisfied. Infact, note that postulate 3 is automatically satisfied since we are dealing with a counting process. And, postulate 4 is there in the form of condition 1. Thus, a counting process satisfying conditions 1-3 is a Poisson process.

From now onwards, therefore, we shall take conditions 1-3 to define a Poisson process.

Example 4: Let us, again, consider a Poisson process $\{X(t), t \geq, 0\}$ of parameter λ.

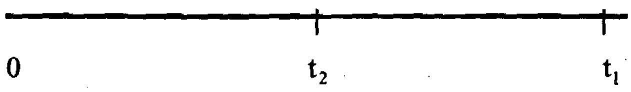
Solution: For $k = 1, 2, \ldots, h$ let T_k be the time between the $(k-1)$ th and k-th event, then $\{T_k : k \geq 1\}$ is the sequence of "inter-arrival times" in the Poisson process. Note that for $n = 1, 2, \ldots$, S_n time $T_1 + T_2 + \cdots + T_n$ is the time instant of occurrence of the nth event. Note that by Eqn. (22), the probability that the waiting time T_1 for the first occurrence of the of event will exceed t turns out to be

$$P(T_1 \geq t) = P(X(t) = 0) = P_0(t) = e^{-\lambda t}$$

which means that T_1 is distributed exponentially with mean $\frac{1}{\lambda}$. Notice, then, that roughly speaking, λ is the rate of occurrence of the event. Again, by independence of occurrence of events over disjoint intervals, by postulate 2, for $0 \leq t_1 \leq t_2 < \infty$ we have

$$\begin{aligned} P(T_1 \leq t_1, S_2 > t_2) &= P(X(t_1) \geq 1, \text{ second occurrence of event after } t_2) \\ &= P(X(t_1) = 1, X(t_2) - X(t_1) = 0) \\ &= P(X(t_1) = 1)P(X(t_2 - t_1) = 0) \\ &= (\lambda t_1)e^{-\lambda t_1} e^{-\lambda(t_2 - t_1)} = (\lambda t_1)e^{-\lambda t_2}. \end{aligned}$$

Consider the case, $0 \leq t_2 \leq t_1 < \infty$, i.e.,



$$\begin{aligned} P(T_1 < t_1, S_2 > t_2) &= P(X(t_1) \geq 1, \text{ second event after } t_2) \\ &= P(\text{ both first and second occurrence of event after } t_2 \text{, but the first event before } t_1) \\ &\quad + P(\text{first occurrence of event before } t_2 \text{, and the second after } t_2) \\ &= P(X(t_2) = 0, X(t_1) - X(t_2) \geq 1) + P(X(t_2) = 1) \\ &= e^{-\lambda t_2}(1 - e^{-\lambda(t_1 - t_2)}) + (\lambda t_2)e^{-\lambda t_2}. \end{aligned}$$

Thus, the joint probability density function of T_1 and S_2 is given by

$$-\frac{\delta}{\delta t_1 \, \delta t_2} P(T_1 < t_1, S_2 > t_2) = \begin{cases} \lambda^2 e^{-\lambda t_2} \text{ for } 0 \leq t_1 \leq t_2 < \infty \\ 0 \text{ for } \qquad 0 \leq t_2 \leq t_1 < \infty \end{cases}$$

Now, by transforming as $X = T_1$ and $Y = S_2 - T_1$, we obtain the joint probability density of T_1 and T_2, which is given by

$$\lambda^2 e^{-\lambda(x+y)}, \ x \geq 0 \text{ and } y \geq 0.$$

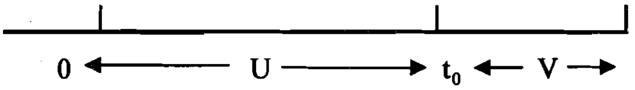
It is, thus, seen that the inter-arrival waiting times T_1 and T_2 are i.i.d with each exponentially distributed mean λ^{-1}.

In Example 3 we showed that the first two inter-arrival times are i.i.d., with each having exponential distribution. Proceeding in the same manner, one can prove a more general property of Poisson process as stated in the theorem given below.

Theorem 2: If $\{X(t): t \geq 0\}$ is a Poisson process with rate λ, then the inter-arrival times $T_1, T_2, \ldots$ are independently distributed each having exponential distribution with mean λ^{-1}. Conversely, suppose $\{T_n : n \geq 1\}$ is a sequence of independently and identically distributed random variables each having the exponential distribution with mean λ^{-1}; define $S_0 = 0$ and for $n = 1, 2, \ldots, S_n = T_1 + T_2 + \cdots + T_n$. For $t \geq 0$, let $X(t) := \max\{n \geq 0 : S_n \leq t\}$. Then $\{X(t): t \geq 0\}$ is a Poisson process.

The above theorem, which we will not prove here, gives a simpler insight into the structure and profile of a Poisson process.

We will now discuss one more important property of the Poisson process. Towards that, visualize a situation where events are occurring in the Poisson fashion over a time scale. For instance, the event can be one of arrival of a public bus at a bus stop. Suppose, you are waiting for a bus at a bus stop and you are interested in the time you may have to wait now, or the time when the last bus went past the bus stop before your arrival. Let t_0 is the time when you arrived at the bus stop. The last bus went past U units of time ago and V is the time that you have to wait to get the next bus. Pictorially,



Thus, in respect of the time point t_0, U is the time since the last occurrence of the event, and V is the waiting time for the next occurrence of the event since t_0. In the context of Reliability Theory which deals with event-like failure of systems, the random variable of interest is the life (i.e., the time to failure) of a system, then U is often referred to as the current life (at the present time t_0) and V is called the excess (to t_0) life. For $0 < u < t_0$ and $v > 0$, the joint distribution function of the random variables U and V can be computed as

$$\begin{aligned}
&P(U \geq u, V \geq v) \\
&= P(X(t_0) - X(t_0 - u) = 0 = X(t_0 + v) - X(t_0)\} \\
&= P\{X(t_0) - X(t_0 - u) = 0\}\, P\{X(t_0 + v) - X(t_0) = 0\} \\
&= e^{-\lambda u} e^{-\lambda v} = e^{-\lambda(u+v)},\ 0 < u < t_0,\ v > 0
\end{aligned}$$

Also, $P(U \geq u) = P(U \geq u, V \geq 0) = e^{-\lambda u}$ for $0 < u < t_0$,
$\qquad = 0$ for $u \geq t_0$

and, $P(V \geq v) = e^{-\lambda v}$, $v \geq 0$.

Thus, U and V are independently distributed.

We will see a few other interesting results through the exercises given below.

E1) For a Poisson process $\{X(t) : t \geq 0\}$, find the conditional probability that there are k events in t units of time, given that there are $k + n$ events in $t + s$ units of time.

E2) For a Poisson process $\{X(t) : t \geq 0\}$, let T_m be the random time to the m-th event. Find its distribution. Also, for $n \geq 1$, obtain $P\{T_1 \leq t \mid X(s + t) = n\}$.

E3) Let the Poisson process $\{X(t) : T \geq 0\}$ with intensity α be independent of another such process $\{Y(t) : t \geq 0\}$ with intensity β. Then $\{Z(t) = X(t) + Y(t) : t \geq 0\}$ is also a Poisson process with intensity $\alpha + \beta$.

E4) Suppose that from a bus stand, one person has the option of traveling to a given destination either by Bus No. 2 or Bus No. 9. Bus No. 2 arrives at the stand following a Poisson process with intensity 6 per hour and Bus No. 9 arrives also following a Poisson process with intensity 4 per hour.

a) Let X denotes the extent of time in hours a person has been late in arriving at the bus stop since the last bus went by. What is the distribution of X?

b) Let W denotes the extent of time in hours the person has to wait for a bus. What is the probability that this person has to wait for more than 10 minutes?

E5) For a Poisson process $\{X(t): t \geq 0\}$, define another process $N(t) = X(t+s_0) - X(t)$, where $s_0 > 0$ is a fixed constant. Is $N(t)$ also a Poisson process?

E6) Suppose a system breaks down from time to time because of the failure of one of two components – component A and component B. The number of system breakdowns follow the Poisson process with parameter λ; also, the probability that a breakdown is caused by the failure of component A is π, so that the probability of a breakdown being caused by the failure of component B is $(1-\pi)$. Show that the number of system breakdowns owing to the failure of component A over time is a Poisson process with parameter $\lambda\pi$.

E7) Suppose in a given society, there are N individuals; moreover the population produces no new individuals, i.e., has birth rate zero, but adds to the population through immigration which takes place as a Poisson process with parameter λ. Let $X(t)$ denotes the size of the population at time t and denote, as usual, by $p_k(t) = P(X(t) = k)$. Develop a suitable differential equation to determine the probability generating function of $X(t)$, and discuss what happens as $t \to \infty$. Also, calculate $EX(t)$.

E8) The time to failure of a certain variety of machine is distributed exponentially with mean $\frac{1}{\alpha}$; the time a mechanic would require, then, to restart the machine is also exponentially distributed with mean $\frac{1}{\beta}$. The failure and repair times are independent of each other. Consider two such machines and two mechanics so that each of the machines can get the immediate attention of a mechanic whenever required. The two machines are put into operation at the same time and let $X(t)$ denotes the number of machines out of action at time t. For $k = 0, 1$, let $p_k(t) = P(x(t) = k)$; show that $p_0(t) = \frac{\beta}{\alpha+\beta}(1-e^{-(\alpha+\beta)})$. Obtain the limiting distribution of $X(t)$ as $t \to \infty$ and show that it has a mean $\frac{2\pi}{(1+\pi)}$, where $\pi = \frac{\alpha}{\beta}$.

Now, we summarise what we have learnt in this unit.

5.5 SUMMARY

In this unit, we covered the following:

1) We learnt a number of concepts that are, in some sense realistic and, hence, are targeted towards modeling processes that we come across in our daily dealings with life.

2) In the previous unit, we dealt with the number of additions to the population through births over generations. Thus, time (generations) was counted on a discrete scale; thus, the first generalization over that concept that one thinks of is to allow time to vary over a continuous scale. Now, more in tune with reality the population is viewed to be growing, or rather changing, over a continuous time scale. We, however, desist at this stage, from even conceiving the process $X(t)$ to be changing continuously itself.

3) We have introduced the Markov property of processes, which, in plain words, is that the future development of a process would depend on the recent-most past.

4) The definition and interesting properties of a Poisson process are discussed, which essentially models occurrences of events that are kind of 'rare' (the probability of occurrence of an event in a time interval is proportional to the length of the said interval, irrespective of its location on the time scale). Obviously, the next behaviour of interest concerning a process has been to study the growth, or otherwise, of a process that is typified by the occurrences of two kinds of events counteracting the effect of each other on the overall change, like births and deaths. Such a process with death rate zero leads to a birth process which is, in another sense, a generalization of the Poisson process.

5.6 SOLUTIONS/ANSWERS

E1) We need to compute $P\{X(t)=k \mid X(t+s)=k+n\}$. Notice that by definition of conditional probability,

$$\begin{aligned}
&P\{X(t)=k \mid X(t+s)=k+n\}\\
&=P\{X(t)=k,\ X(t+s)=k+n\}/P\{X(t+s)=k+n\}\\
&=P(X(t)=k,\ X(t+s)-X(t)=n\}/P(X(t+s)=k+n\}\\
&=P\{X(t)=k\}\ P\{X(t+s)-X(t)=n\}/P(X(t+s)=k+n),
\end{aligned}$$

since in a Poisson process, the number of events in disjointed intervals are independent.

$$\begin{aligned}
&=\{e^{-\lambda t}(\lambda t)^k/k!\}\ \{e^{-\lambda s}(\lambda s)^n\}/\{e^{-\lambda(t+s)}[\lambda(t+s)]^{k+n}/(k+n)!\}\\
&=\{(k+n)!/k!\ n!\}\ \{t/(t+s)\}^k\ \{s/(t+s)\}^n.
\end{aligned}$$

E2) We may represent S_m as $S_m=T_1+T_2+\cdots+T_m$, where T_k is the time between the $(k-1)$-th and k-th event, $1\le k\le m$. From Theorem 2, we know that the inter-arrival times $T_1, T_2, \ldots$ are independently distributed each having an exponential distribution with mean λ^{-1}.

Clearly,

$$\begin{aligned}
&P(S_m\le t)\\
&=P\{X(t)\ge m\}\\
&=\sum_{k=m}^{\infty} e^{-\lambda t}(\lambda t)^k/k!,
\end{aligned}$$

so that the pdf of S_m, which is the time to the m-th event is

$$\begin{aligned}
&\{d/dt\}\ P(S_m\le t)\\
&=-\lambda e^{-\lambda t}\sum_{k=m}^{\infty}(\lambda t)^k/k!+e^{-\lambda t}\lambda\sum_{k=m}^{\infty}k(\lambda t)^{k-1}/k!\\
&=-\lambda e^{-\lambda t}\sum_{k=m}^{\infty}(\lambda t)^k/k!+\lambda e^{-\lambda t}\sum_{k=m-1}^{\infty}(\lambda t)^k/k!\\
&=e^{-\lambda t}\lambda^m t^{m-1}/(m-1)!,\ t\ge 0.
\end{aligned}$$

Now,

$$\begin{aligned}
&P\{T_1\le t \mid X(s+t)=n\}\\
&=P\{X(t)\ge 1 \mid X(t+s)=n\}\\
&=P\{X(t)\ge 1,\ X(t+s)=n\}/P\{X(t+s)=n\}\\
&=\sum_{k=1}^{\infty} P\{X(t)=k,\ X(t+S)=n\}/P\{X(t+s)=n\}
\end{aligned}$$

$$=\sum_{k=1}^{\infty} P\{X(t)=k,\ X(t+S)-X(t)=n-k\}/P\{X(t+s)=n\}$$

$$=\sum_{k=1}^{\infty} P\{X(t)=k\}\ P\{X(t+S)-X(t)=n-k\}/P\{X(t+s)=n\}$$

Since, in a Poisson process, the number of events in disjoint intervals are independent.

$$=\sum_{k=1}^{\infty} \{e^{-\lambda t}(\lambda t)^k/k!\}\ \{e^{-\lambda s}(\lambda s)^{n-k}/(n-k)!\}$$

$$x\{n!/\{e^{-\lambda(t+s)}[\lambda(t+s)]^n\}$$

$$=\sum_{k=1}^{\infty} \frac{n!}{k!(n-k)!}\left(\frac{t}{t+s}\right)^k\left(\frac{s}{t+s}\right)^{(n-k)}$$

$$=\left\{\left(\frac{t}{t+s}\right)+\left(\frac{s}{t+s}\right)\right\}^n-\left(\frac{s}{t+s}\right)^n$$

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

E3) We need to argue that

(i) $Z(0)=0$

(ii) $\{Z(t):t>\}$ has independent increments

(iii) for any $h>0,\ Z(t+h)-Z(t)$, which is the number of events (birth) in the interval $(t,\ t+h]$ of length h, follows the Poisson distribution with mean $(\alpha+\beta)h$.

Since $Z(t)=0\equiv X(0)=0,\ Y(0)=0$, (i) is obvious since $X(t)$ and $Y(t)$ themselves are Poisson processes. Again, because of the same reason, (ii) easily follows. Now, $Z(t+h)-Z(t)=\{X(t+h)-X(t)\}+\{Y(t+h)\}$ and each of the components on the right hand side is a Poisson random variable and they are independent as $X(t)$ and $Y(t)$ are also independent. Hence, their sum, i.e. $Z(t+h)-z(t)$ is a Poisson random variable for all $t\geq 0,\ h\geq 0$. This establishes that $\{Z(t):t\geq 0\}$ is a Poisson process.

E4)* Let $X(t)$ denotes the number of Bus No.2 that pass through the bus stand by time t and similarly let, $Y(t)$ the number of Bus No.2 that pass through the same bus stand; thus, since either of the two kinds of buses are fine for the person, the process of interest for the person would be $Z(t)=X(t)+Y(t)$. Again since both $X(t)$ and $Y(t)$ are Poisson processes, so is $Z(t)$ with intensity 10, assuming $X(t)$ and $Y(t)$ to be independent. As such, from theory,

$$P(X\geq u)=e^{-\lambda u} \text{ for } 0<u<t_0,$$
$$=0 \qquad \text{for } u\geq t_0$$

where t_0 is the time that the person arrived at the bus stand. Also, from theory, W has an exponential distribution with intensity 10. Accordingly, $P(W\geq 10)=e^{-10/10}=e^{-1}$.

E5) Let us try $P\{N(t)=k\}=P\{X(t+s_0)-X(t)=k\}=e^{-\lambda s_0}(\lambda s_0)^k/k!\}$ for all t. The right hand side is free of t!. Thus, $N(t)$ cannot be construed as a Poisson process.
Let us now look at $U(t).P\{U(t)=k\}=P\{X(t+s_0)-X(s_0)=k\}$

$$=e^{-\lambda t}(\lambda t)^k/k!.$$

Also, by definition $U(0)=0$. Finally, since $X(t)$ has independent increments, so does $U(t)$ as well. Hence, $U(t)$ is a Poisson process.

E7) Let $Y(t)$ denotes the number of persons added through immigration to the population so that $X(t)=Y(t)+N$. Now, for $k\geq N$, since $Y(t)$ is a Poisson process with parameter λ,

$$\begin{aligned}
p_k(t+h)=P(X(t+h)=k &= \sum_{j=0}^{k} P\{X(t+h)-X(t)=k-j,\ X(t)=j\} \\
&= P\{Y(t+h)-Y(t)=0,\ k-N\}+ \\
&\quad + P\{Y(t+h)-Y(t)=1,\ Y(t)=k-N-1\}+0(h) \\
&= P\{Y(t+h)-Y(t)=0\}\ P\{Y(t)=k-N\}+ \\
&\quad + P\{Y(t+h)-Y(t)=1\}\ P\{Y(t)=k-N-1\}+0(h) \\
&= P\{Y(t+h)-Y(t)=0\}\ P\{X(t)=k\}+ \\
&\quad + P\{Y(t+h)-Y(t)=1\}\ P\{X(t)=k-1\}+0(h) \\
&= e^{-\lambda h}\,p_k+\lambda h\, e^{-\lambda h}\,p_{k-1}(t)+0(h) \\
&= \{1-\lambda h+0(h)\}\ p_k(t) \\
&\quad +\lambda h\{1-\lambda h+0(h)\}\,p_{k-1}(t)+0(h), \text{ so that}
\end{aligned}$$

$\{p_k(t+h)-p_k(t)\}/h=-\lambda p_k(t)+\lambda p_{k-1}(t)+0(1)$, and hence, as $h\to\infty$,

$$p_k'(t)=-\lambda p_k(t)+\lambda p_{k-1}(t),\ k\geq N. \qquad (5.7.1)$$

For $k=N$, $p_N'(t)=-\lambda p_N(t)+\lambda p_{N-1}(t)=-\lambda p_N(t)$, since $p_{N-1}(t)=0$, as the population can never be smaller than the number of original individuals.
Solving the differential equation, $p_N'(t)=-\lambda p_N(t)$, we find $e^{\lambda t}p_N(t)=C$, a constant. But, $p_N(0)=1$, since, again the original population size is N, so that $C=1$ and as such $p_N(t)=e^{-\lambda t}$. Now from Eqn. (5.7.1),

$p_{N+1}'(t)+\lambda p_{N+1}(t)=\lambda p_N(t)=\lambda e^{-\lambda t}$, i.e. $\dfrac{\partial}{\partial t}\{e^{\lambda t}\,p_{N+1}(t)\}=\lambda$, i.e.

$e^{\lambda t}\,p_{N+1}(t)=\lambda t+\text{constant}$.

Again, $p_{N+1}(0)=0$, so that the Constant = 0 and as such,

$$p_{N+1}(t)=\lambda t\,e^{-\lambda t};$$

solving successively, $p_{N+k}(t)=(\lambda t)^k\,e^{-\lambda t}/k!,\ k=0,1,2,\ldots$

---x---

UNIT 6 CONTINUOUS TIME MARKOV PROCESSES-II

Structure

6.1 INTRODUCTION

In the last unit, we considered the pure birth process as an example of a continuous time Markov process where the births occur randomly with a given rate in any small time interval. In Sec. 6.2, we will once again look at pure birth processes. Analogous to the pure birth process where the state of the system changes only on the birth of another individual, we shall consider a pure death process in Sec. 6.3, where deaths occur at random times with a prescribed rate. For example, such a situation can occur in a system of m components which fail independently of one another. Then, the process which counts the number of components which are working at time t will be a pure death process. One can also visualise a system (e.g., the population of a certain species) where both births and deaths occur at random times. In Sec. 6.4, we will introduce and study such processes and see several examples. We will also study the steady state distribution of the processes in Sec. 6.5.

Objectives

After studying this unit, you should be able to:

- identify pure birth, pure death and birth and death processes;
- find the transition probabilities for these processes;
- write down the infinitesimal generator of these processes;
- find the steady state probabilities for these processes.

6.2 PURE BIRTH PROCESS

Recall that the pure birth process was introduced in Sec. 5.4 as an example of a continuous time branching process. In particular, it is a Markov process on the state space of non-negative integers where the only transitions that are possible are due to births that take place at random times. In other words, it is a collection of random variables

$$\{X_t : t \geq 0\}$$

where the transition (jump) from X_t depends only on the (current) value of X_t, and not on the past of the process (Markov property). Further, if $X_t = k$, a birth occurs at random time in small time interval $(t, t+h)$ with some rate λ_k which, in general, might depend on k. It is assumed that the chance of more than one birth occurring simultaneously in infinitesimal interval is negligible. These conditions can be written compactly, as was done in the previous unit.

Let $P_{i,j}(t)$ denotes the transition probability

$$P_{i,j}(t) = P(X_t = j \mid X_0 = i) = P(X_{s+t} = j \mid X_s = i) \qquad (1)$$

Then, for $k \geq 0$,

(b1) $P_{k,\,k+1}(h) = \lambda_k h + g_1(h)$

(b2) $P_{k,\,k}(h) = 1 - \lambda_k h + g_0(h)$

(b3) $P_{k,\,k-j}(h) = g_j(h)$ for $j = -1, \pm 2, \pm 3, \ldots$

(b4) $X_0 = 0$

where $g_j(h)$ $j \geq 0$ are of order $0(h)$ as $h \downarrow 0$.

We will see several example of pure birth processes in this section.

Example 1 (Poisson Process): The simplest (and perhaps the most important) example is when the birth rates λ_k do not depend on k. If we denote the common rate by λ, the corresponding pure birth process is called the Poisson process with parameter λ. Such processes arise in many real life situations where the "births" occur at a constant rate irrespective of the elapsed time as well as the number of "births" that have already occurred. Here are a few such examples.

(a) : Let X_t denotes the number of fish caught from a lake by an individual in the time interval $[0, t]$. Under the assumption that the number of fish in the lake is quite large, the rate at which the fish will be hooked can be considered to be a constant and, hence, X_t may be considered to be a Poisson process, where the catch of each fish is a "birth". Since the fish being caught in the fishing hook at any instant is as likely as any other instant, the births do not depend on time t

(b) : The fishing example above is a special case of general phenomena observed in queueing models. If X_t denotes the number of arrivals into the queue in the time interval $[0, t]$, then X_t is a Poisson process if arrivals follow the assumptions underlying such process.

(c) : Consider a radioactive particle which decays with time. A Geiger counter detects these disintegrations. Let X_t denotes the number of disintegrations detected by the counter in time interval $[0, t]$. Then, X_t is a Poisson process since, in most cases, the half-life of the radioactive substance, or the time it takes for the substance to decay to exactly half its mass, is large and, hence, the rate of decay can be considered to be a constant. Needless to say, in this example, the "births" considered are decays of a radioactive particle.

Example 2 (Yule Process): This was also introduced in the previous unit by putting $\lambda_k = k\lambda$. A real life situation where such a process arises is in the population of amoebae. Each amoeba lives for a random amount of time after which it splits in two (so that the population size goes up by one). Let X_t denotes the population of the colony at time t. If all amoebae live and split independently of one another, then it can be shown that X_t is a Yule process.

Example 3 (Yule Process with Immigration): Suppose that in the amoebae colony in Example 2, it is possible that amoebae also come from outside. Suppose that the stream of amoebae which enter from outside is like a Poisson process with parameter λ. All the amoebae in the colony continue to split as before. Then the population process X_t will be a pure birth process where the birth rates $\lambda_k = k\lambda + \theta$.

Example 4 (Maintenance Process): Consider a machine whose lifetime is an exponential random variable with parameter λ_0. On its failure, suppose that it is repaired instantaneously. The lifetime of the repaired machine is now exponential λ_1. This process is repeated, so that after the n^{th} repair, the lifetime of the machine is exponential with parameter λ_n. If X_t denotes the number of repairs carried out in time interval $[0, t]$, then $\{X_t\}$ is a pure birth process with birth rate λ_n depending on the number of births n so far.

In the previous unit, the infinitesimal generator $\mathbf{Q} = ((Q_{i,j}))$ was introduced (see section 5.2) where $Q_{i,j}$ are the infinitesimal transition rates defined by

$$Q_{i,j} = \lim_{h \to 0} \frac{P_{i,j}(h) - P_{i,j}(0)}{h}$$

where

$$P_{i,j}(0) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Using the postulates **(b1)** – **(b4)** on the transition probabilities of a pure birth process one can easily calculate the infinitesimal transition rates. We leave this as an exercise.

Now try the following exercise.

E1) Show that for a pure birth process infinitesimal transition rates are given by

$$Q_{i,j} = \begin{cases} -\lambda_i & \text{if } j = i \\ \lambda_i & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

Thus, the infinitesimal generator $\mathbf{Q}$ is the matrix of the form

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ 0 & -\lambda_1 & \lambda_1 & 0 & \cdots \\ 0 & 0 & -\lambda_2 & \lambda_2 & \cdots \\ 0 & 0 & 0 & -\lambda_3 & \lambda_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

Now, we look at "waiting times" associated with pure birth processes.

Definition 1: The waiting times between births is a sequence of random variables $\{T_k\}$, where T_k denotes the time between the k^{th} and the $(k+1)^{th}$ birth.
In the previous unit, it was shown that (see Eqn. (19))

$$P_0(t) = P(X_t = 0) = \exp(-\lambda_0 t)$$

which, in turn, implied that T_0 has an exponential distribution with parameter λ_0 (or mean $1/\lambda_0$). We can use Eqn. (20) of the previous unit to show inductively that T_k has exponential distribution with parameter λ_k and it is also independent of $T_0, \ldots, T_{k-1}$. Here, T_k denotes the amount of time that the process spends in state k, after which it jumps by 1 to state $k+1$. Thus, pure birth process can be called a pure jump process.

The pure birth process can also be defined by its waiting time distributions. Let T_k be an independent sequence of exponential random variables with parameters λ_k respectively. Then, define

$$X_t = k \text{ if } \sum_{i=0}^{k} T_i \le t < \sum_{i=0}^{k+1} T_i \qquad (2)$$

Then, it can be shown that X_t is a pure birth process in the sense that it satisfies the defining postulates **(b1) – (b4)**. We summarise the above discussion in the following theorem which we state without proof. (The proof in a special case is given in Unit 7).

Theorem 1: Let $\{X_t, t \ge 0\}$ be an increasing jump process, and $\{T_k\}$ be a sequence of random variables such that Eqn. (2) holds. Then, $\{X_t\}$ is a pure birth process with birth rates $\{\lambda_k\}$ if and only if T_k's are independent random variables and T_k has an exponential distribution with parameter λ_k.

E2) In Example 2, assume that the lifetime of each amoeba is exponential λ and is independent of the lifetimes of all other amoeba. Show that the population process X_t defined there is a Yule process.

E3) Show that the process X_t of Example 3 is a pure birth process, and find the birth rates under the assumption of E2) above.

So far, we have discussed the pure birth process and the Poisson process. In the following section, we shall discuss a pure death process.

6.3 PURE DEATH PROCESS

In the previous section we looked at pure birth processes, where the process stays in a state k for an exponential amount of time, T_k, after which it jumps. The only jumps which are allowed are of size $+1$, so that when the jump occurs the state changes from k to $k+1$. We can, also consider pure jump processes where the jumps allowed are of size -1. So that a process can jump from k, only to $k-1$. Such a process will be called a pure death process.

Example 5 (System with Parallel Components): Consider a system with built-in redundancy, so that n identical parts are connected in parallel. The system works as long as at least one of the parts work. We assume that the lifetimes of the components are independent exponential random variables with parameter, μ. Let X_t denotes the number of parts out of n which are working at time t. Then X_t takes values in the state space $\{0, 1, 2, \ldots, n\}$. If $X_t = i$ at some time, t, then the process will jump to state $i-1$, if any one of the i components fails. During infinitesimal small time interval $(t, t+h)$. If the waiting time in state i, is denoted by T_i, then the distribution of T_i is that of the minimum of i independent and identically distributed $\exp(\mu)$ random variables. Thus, T_i is $\exp(i\mu)$ random variable. The process $\{X_t : t \ge 0\}$ is then a pure death process.

Note that for a pure death process taking values in the set of non-negative integers, the state 0 becomes an absorbing state, since no transition is possible from 0. This is the case in the above example as well. As we did for the pure birth process, we can write down the postulates for a pure death process.

We will continue the same notation for transition probabilities.

(d1) $P_{k,k-1}(h) = \mu_k h + g_1(h)$, for $k \ge 1$

(d2) $P_{k,k}(h) = 1 - \mu_k h + g_0(h)$, for $k \ge 1$

(d3) $P_{k,\,k+j}(h) = g_j(h)$ for $j = 1, \pm 2, \pm 3, \ldots$

(d4) $P_{0,\,0}(h) = 1$

where $g_j(h)$ are all of order $0(h)$ as $h \downarrow 0$. Note the postulate **(d4)**. Of course, it can also be incorporated in **(d1)** and **(d2)** by requiring that $\mu_0 = 0$. The sequence $\{\mu_n : n \geq 1\}$ are called the death rates of the process.

Note that a pure death process is also a pure jump process and, hence, the results on waiting times (viz., Theorem 1) remains valid (except, of course for T_0). Just as we did for the pure death process, we can calculate the infinitesimal rates $Q_{i,\,j}$ and the infinitesimal generator **Q** of the system. We leave this as an exercise.

Before, we proceed further, try this exercise.

E4) Consider a pure death process X_t with death rates $\{\mu_n\}$. Find the infinitesimal generator **Q**.

In this section, we shall study a birth and death process.

6.4 BIRTH AND DEATH PROCESSES

In the earlier two sections, we looked at jump processes where all jumps of the process are $+1$, or all jumps are -1. It is conceivable that we encounter processes where both births and deaths can occur, i.e., jumps of $+1$ as well as -1 are allowed. Indeed this is the case in many real life situations. We start with a few examples.

Example 6 (Single Server Queueing Models): Consider a service facility where customers arrive at random times as a Poisson process with parameter λ. So that the lengths time intervals between successive arrival of customers is an i.i.d. sequence of exponential (λ) random variables. There is only one service counter so that only one customer is attended to at a given time. All the other arrivals (if any) form a queue for getting service. Suppose further that the service time for each customer is also exponentially distributed with parameter μ. Let X_t denotes the queue length at time t. Then, X_t either jumps up by one if a new customer arrives in the queue, or decreases by one if a customer leaves the service counter, after being attended to. Thus, X_t is a birth and death process with birth parameter $\lambda_n \equiv \lambda$ and death parameter $\mu_n \equiv \mu$.

Example 7 (Linear Growth Model): Consider a colony of individuals whose lifetimes are independent and identically distributed exponential random variables with parameter μ. Each individual either gives birth to another individual of the same type, independently of one another, or dies at random times which are distributed as exponential random variables with parameter λ. Let X_t denotes the population size at time t. Then, $\{X_t : t \geq 0\}$ is a birth and death process.

Under the assumptions mentioned above, and proceeding, as we did for the Yule process, we can identify the birth and death rates for the linear growth model. We leave this as an exercise.

Example 8: Consider a system with N identical machines which work (or fail) independently of one another. Each machine fails at an exponential (μ) time. Once the machine has failed it is repaired. The repairs take place independently of everything else (for example, each machine may have its own repairman, and then it does not matter as all the machines are down at a certain time), and the repair time for each machine has an exponential distribution with parameter λ. Let X_t denotes the number of machines working at time t. Then, X_t is a birth and death process (with death for the cited example) with parameters $\mu_i = i\mu$ and birth parameters $\lambda_i = (N-i)\lambda$ for $i = 0, 1, \ldots, N$.

Now, try the following exercises.

E5) Find λ_n and μ_n in Example 7 above.
At this time, let us look at the defining postulates for a birth and death process. We continue to use the same notation for transition probabilities.

(bd1) $P_{k,k-1}(h) = \mu_k h + g_1(h)$, for $k \geq 1$

(bd2) $P_{k,k+1}(h) = \lambda_k h + g_2(h)$, for $k \geq 0$

(bd3) $P_{k,k}(h) = 1-(\lambda_k + \mu_k)h + g_3(h)$, for $k \geq 0$

(bd4) $P_{i,j}(0) = \delta_{i,j}$

(bd5) $\mu_0 = 0$

where $g_i(h)$ is of the order $0(h)$ as $h \downarrow 0, i = 1, 2, 3$. Using these postulates, it is now easy to find the infinitesimal generator of the birth and death process.

E6) Show that for the birth and death process with birth rates $\{\lambda_k\}$ and death rates $\{\mu_k\}$, the infinitesimal generator, $\mathbf{Q}$, is given by

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

As we did in the previous unit for the Poisson process, now we can write the differential equations of the birth and death process.

Lemma 1: Let $\{X_t : t \geq 0\}$ be a birth and death process, with death parameters $\{\mu_n\}$, and birth parameters $\{\lambda_n\}$ respectively, so that $P_{i,j}(t)$, defined as by (bd1) to (bd5) are the transition probabilities. Then

1. the backward Kolmogorov's differential equations are given by
$$P'_{0,j}(t) = -\lambda_0 P_{0,j}(t) + \lambda_0 P_{1,j}(t)$$
$$P'_{i,j}(t) = \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i)P_{i,j}(t) + \lambda_i P_{i+1,j}(t), \quad i \geq 1 \text{ for } j \geq 0 \qquad (3)$$
with $P_{i,j}(0) = \delta_{ij}$.

2. the forward Kolmogorov's differential equations are given by
$$P'_{i,o}(t) = -\lambda_0 P_{i,0}(t) + \mu_1 P_{i,1}(t)$$
$$P'_{i,j}(t) = \lambda_{j-1} P_{i,j-1}(t) - \lambda_j + \mu_j)P_{i,j}(t) + \mu_{j+1} P_{i,j+1}(t), \quad j \geq 1 \text{ for } j \geq 0 \qquad (4)$$
with $P_{i,j}(0) = \delta_{ij}$.

Proof: The proof follows directly from Eqn. (5) and Eqn. (6) of the previous unit, and the form of the infinitesimal generator **Q** derived in E6).

We will use the above result in the next section to compute the steady state probabilities for a birth and death process.

6.5 STEADY STATE PROBABILITIES

We ended the previous section with lemma 1 which gives the Kolmogorov backward and forward differential equations ((3) and (4), respectively). These are nothing but difference-differential equations involving the transition probabilities $P_{i,j}(t)$ and thus they tell us as to how these transition probability functions evolve with time. In this section, we will look at the long term behaviour of these functions. We note that if $\lambda_0 = 0$, the state 0 is absorbing, and then the state space is not irreducible, so we assume that $\lambda_0 > 0$. We need the following result which we state without proof.

Theorem 2: For a birth and death process with the transition probability function given by $P_{i,j}(t)$, as in E5).

$$\lim_{t\to\infty} P_{i,j}(t) = \Pi_j \tag{5}$$

exists. And, it is not dependent on initial state i.

Note that the limit depends only on j, and the initial state i is "forgotten". We will use the above result to find the Π_j's in terms of the birth and death parameters $\{\lambda_n\}$ and $\{\mu_n\}$, respectively. The question that we need to answer now is: under what conditions do limiting probabilites form a probability distribution. First, we have the following definition.

Definition 2: Suppose Π_j are as in Eqn. (5). If

$$\sum_j \Pi_j = 1 \tag{6}$$

the sequence $\{\Pi_j\}$ is called an **equilibrium,** or **steady state distribution** of the birth and death process.

To study this sequence in detail, we recall the Chapman Kolmogorov equations. (See Eqn. (2) of Unit 5.) For any states k, j, and any $t, s \geq 0$, we have

$$P_{k,j}(s+t) = \sum_{i=0}^{\infty} P_{k,i}(s)\, P_{i,j}(t)$$

Letting $s \to \infty$ and using Eqn. (5) (and, assuming that we can interchange limit and sum), we get

$$\Pi_j = \sum_{i=0}^{\infty} \Pi_i P_{i,j}(t) \tag{7}$$

The above identity can be interpreted as follows. Suppose steady state distribution exists for a continuous time Markov Chain, so that $\{\Pi_j\}$ is a probability distribution. Suppose that for every i, Π_i is taken as the probability that the process is in state i initially. Then, the probability that the process is in state j after any given time t (given by the right hand side of Eqn. (7)) is Π_j. Hence, the probability distribution of the state of the system at anytime in future is same as the initial distribution of the sequence $\{\Pi_j\}$ satisfying Eqn. (7) subject to (6) stationary distribution of the system. Thus, steady state distribution if it exists is also the stationary distribution.

The above can be restated as follows. If the process starts in state j and t is large, then

$$P_{j,k}(t) = \Pi_k \text{ for all } j, k$$

This, in turn, implies that for all j and k

$$P'_{j,k}(t) = 0.$$

Putting these values in the Kolmogorov forward Eqn. (4), we get

$$0 = -\lambda_0 \Pi_0 + \mu_1 \Pi_1$$

$$0 = \lambda_{j-1}\Pi_{j-1} - (\lambda_j + \mu_j)\Pi_j + \mu_{j+1}\Pi_{j+1}, \quad j \geq 1 \tag{8}$$

These equations can also be interpreted as follows. In the steady state, the rate at which the (birth and death) process enters state j is equal to the rate at which the process leaves state j. For example, when $j = 0$, the process will leave state 0 and jump to state 1, with rate $\lambda_0 \Pi_0$. At the same time, it will enter state 0 (by jumping from 1, following a death) with rate, $\mu_1 \Pi_1$. Equating the two will give the first of the two equalities in Eqn.(8) above. The second equality can be derived similarly.

Our aim now, is to solve the system of equations in Eqn. (8). We will use induction to first express Π_j in terms of Π_0. The first equation in Eqn. (8) gives

$$\Pi_1 = \frac{\lambda_0}{\mu_1}\Pi_0$$

Now assume that for some $k \geq 1$, we have

$$\Pi_k = \frac{\lambda_0 \lambda_1 \cdots \lambda_k}{\mu_1 \mu_2 \cdots \mu_k}\Pi_0 \tag{9}$$

Then for $k+1$ we get from the second equality of Eqn. (9)

$$\begin{aligned}
\mu_{k+1}\Pi_{k+1} &= (\lambda_k + \mu_k)\frac{\lambda_0 \lambda_1 \cdots \lambda_k}{\mu_1 \cdots \mu_k}\Pi_0 - \lambda_{k-1}\frac{\lambda_0 \lambda_1 \cdots \lambda_{k-2}}{\mu_1 \mu_2 \cdots \mu_{k-1}}\Pi_0 \\
&= \lambda_k \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}\Pi_0 + \left(\mu_k \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} - \lambda_{k-1}\frac{\lambda_0 \lambda_1 \cdots \lambda_{k-2}}{\mu_1 \mu_2 \cdots \mu_{k-1}}\right)\Pi_0 \\
&= \lambda_K \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}\Pi_0.
\end{aligned} \tag{10}$$

Thus, finally, we get

$$\Pi_{K+1} = \frac{\lambda_0 \lambda_1 \cdots \lambda_k}{\mu_1 \mu_2 \cdots \mu_{k+1}}\Pi_0$$

and Eqn. (9) holds for all k.

Finally, to check that Eqn. (6) is satisfied, we need that

$$\Pi_0 + \sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}\Pi_0 = 1$$

which gives

$$\Pi_0 = \left(1 + \sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}\right)^{-1} \tag{11}$$

Note that in the above equation, $\Pi_0 > 0$ if, and only if

$$\sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} < \infty. \tag{12}$$

Thus, Eqn. (12) gives a necessary and sufficient condition in terms of the birth and death rates for the birth and death process to have a stationary distribution. Thus, this is the condition for all states to be positive recurrent.

The above results can be used to do the following exercises.

E7) Consider the single server queueing model of Example 6. When will a stationary distribution exist? Whenever it does exist find Π_k.

Let us look at a couple of examples of birth and death processes, and calculate their steady state probabilities.

Example 9 (Simple Birth and Linear Death Process): Consider a birth and death process where the birth rate is constant $(\lambda_n \equiv \lambda)$, and the death rate is linear so that $\mu_n = n\mu$ for all n. We will investigate the long term behaviour of the process. First, note that

$$\sum_{k=1}^{\infty} \frac{\lambda_0\lambda_1 \ldots \lambda_{k-1}}{\mu_1\mu_2 \ldots \mu_k} = \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k < \infty$$

for all $\lambda, \mu > 0$. Thus, a stationary distribution always exists. Further

$$\Pi_0 = \left(1 + \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\right)^{-1} = \exp\left\{-\frac{\lambda}{\mu}\right\}.$$

Now Eqn. (9) implies

$$\Pi_k = \frac{(\lambda/\mu)^k}{k!} \exp\{-\lambda/\mu\}.$$

Example 10: Consider the birth and death process $\{X_t\}$ of Example 5. Here $\{X_t\}$ denotes the number of machines out of N that are working at time t. As observed earlier X_t can take values in the set $\{0, 1, \ldots, N\}$. Further, for $i = 0, 1, \ldots, N$, we have $\mu_i = i\mu$ and $\lambda_i = (n-i)\lambda$. Putting $\lambda_i = 0$ for $i > N$, we see that

$$\sum_{k=1}^{\infty} \frac{\lambda_0\lambda_1 \ldots \lambda_{k-1}}{\mu_1\mu_2 \ldots \mu_k} = \sum_{k=1}^{\infty} \frac{\lambda_0\lambda_1 \ldots \lambda_{k-1}}{\mu_1\mu_2 \ldots \mu_k} < \infty.$$

Further

$$\Pi_0 = \left(1 + \sum_{k=1}^{N} \frac{(N(n-1)\ldots(N-k+1))\lambda^k}{k!\mu^k}\right)^{-1}$$

$$= \left(\sum_{k=0}^{N}\binom{N}{k}\left(\frac{\lambda}{\mu}\right)^k\right)^{-1}$$

$$= \left(1 + \frac{\lambda}{\mu}\right)^{-N}$$

Thus

$$\Pi_k = \binom{N}{k}\left(\frac{\lambda}{\mu}\right)^k \left(1 + \frac{\lambda}{\mu}\right)^{-N}, \quad k = 0, 1, \ldots, N.$$

We end this unit with an exercise.

E8) Consider a birth and death process with parameters $\lambda_k = \lambda/(k+1)$ and $\mu_k = \mu$ for all $k \geq 0$, where $\lambda, \mu > 0$. Show that a stationary distribution always exists. Find Π_k.

Now, let us end this unit by summarizing the following points.

6.6 SUMMARY

In this unit, we studied the following:

1. We defined a pure death process using postulates about its transition mechanisms. We saw several examples from real life situations where such processes arise. We also learnt to calculate the birth rates, and studied its infinitesimal generator.
2. We also looked at a pure death process and its infinitesimal generator. We also saw some real life examples.
3. We studied the birth and death process via its defining postulates, saw many examples and identified its infinitesimal generator. We also derived the forward and backward Kolmogorov differential equations.
4. We studied the long term behaviour of a birth and death process by looking at its steady state probabilities, and deriving them in terms of $\{\lambda_n\}$ and $\{\mu_n\}$. We also found necessary and sufficient conditions for the steady state probabilities to exist.

6.7 SOLUTIONS/ANSWERS

E1) $$Q_{i,i} = \lim_{h\to 0} \frac{P_{i,i}(h) - P_{i,i}(0)}{h} = \lim_{h\to 0} \frac{1 - \lambda_i h + g_0(h) - 1}{h} = -\lambda_i.$$

$$Q_{i,i+1} = \lim_{h\to 0} \frac{P_{i,i+1}(h) - P_{i,i+1}(0)}{h} = \lim_{h\to 0} \frac{\lambda_i h + g_1(h)}{h} = \lambda_i.$$

E2) We need to use the following result. If $Y_1, Y_2, \ldots, Y_n$ are independent exponential random variables with parameters $\theta_1, \theta_2, \ldots, \theta_n$ respectively, then $Z = \min\{Y_1, Y_2, \ldots, Y_n\}$ is exponential with parameter $\sum_{i=1}^{n} \theta_i$.
In Example 2, the population size jumps from k to $k+1$ as soon as any of the k amoebae split, which is the minimum of k i.i.d. exponential (λ) random variables. Hence, the waiting time in state k is exponential with parameter $k\lambda$, and Theorem 1 now gives the result. Independence of T_k follows from independence of lifetime distributions and memory less property of exponential distribution.

E3) We argue, as in E2, above. The only difference here is that the T_k the minimum of k i.i.d. exponential (λ) and one exponential (θ) random variables. Thus, $T_k \sim \exp(k\lambda + \theta)$.

E4) For $i \geq 1$,

$$Q_{i,i} = \lim_{h\to 0} \frac{P_{i,i}(h) - P_{i,i}(0)}{h} = \lim_{h\to 0} \frac{1 - \mu_i h + g_0(h) - 1}{h} = -\mu_i$$

$$Q_{i,i-1} = \lim_{h\to 0}\frac{P_{i,i-1}(h)-P_{i,i-1}(0)}{h} = \lim_{h\to 0}\frac{\mu_i h + g_1(h) - 1}{h} = \mu_i$$

Further $Q_{0,0} = 0$. Thus,

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \mu_1 & -\mu_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -\mu_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -\mu_3 & 0 \\ 0 & 0 & 0 & \mu_4 & -\mu_4 \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

E5) Note that X_t can take values in the set {0, 1, 2, ...}. Suppose $X_t = k$ for some $k \geq 1$. Then, X_t will jump up by one as soon as any one of the k individuals gives a birth. This random time (as seen in the earlier exercises) is an exponentially distributed random variable, with parameter $k\lambda$. Similarly, X_t will decrease by 1 if any one of the k individuals dies. This will happen at a random time which is exponentially distributed with parameter $k\mu$. Note further that 0 is an absorbing state. Thus, X_t is a birth and death process with birth parameters $\lambda_k = k\lambda$, and death parameters $\mu_k = k\mu$ for $k \geq 0$.

E6) For $i \geq 1$,

$$Q_{i,i-1} = \lim_{h\to 0}\frac{P_{i,i-1}(h)-P_{i,i-1}(0)}{h} = \lim_{h\to 0}\frac{\mu_i h + g_1(h) - 1}{h} = \mu_i$$

For $i \geq 0$,

$$Q_{i,i+1} = \lim_{h\to 0}\frac{P_{i,i+1}(h)-P_{i,i+1}(0)}{h} = \lim_{h\to 0}\frac{\lambda_i h + g_2(h) - 1}{h} = \lambda_i$$

$$Q_{i,i} = \lim_{h\to 0}\frac{P_{i,i}(h)-P_{i,i}(0)}{h} = \lim_{h\to 0}\frac{1-(\lambda_i+\mu_i)h + g_3(h) - 1}{h} = -(\lambda_i+\mu_i).$$

E7) For the birth and death process of E6 the birth and death parameters are identified as $\lambda \equiv_n \lambda$, and $\mu_n \equiv \mu$ respectively, for all n. Then

$$\sum_{k=1}^{\infty}\frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k} = \sum_{l=1}^{\infty}\left(\frac{\lambda}{\mu}\right)^k$$

The right hand side is finite, if and only if, $\lambda < \mu$, and under this condition distribution exists. In this case, using Eqn. (11), we get

$$\Pi_0 = \left(1+\sum_{l=1}^{\infty}\left(\frac{\lambda}{\mu}\right)^k\right)^{-1} = \left(1-\frac{\lambda}{\mu}\right)$$

Now Eqn. (9) implies

$$\Pi_k = \left(\frac{\lambda}{\mu}\right)^k\left(1-\frac{\lambda}{\mu}\right).$$

E8) *To check that a stationary distribution exists*, we consider the expression in Example 12.

$$\sum_{k=1}^{\infty}\frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k} = \sum_{k=1}^{\infty}\frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k < \infty \text{ for all } \lambda, \mu.$$

This proves the existence of a stationary distribution whatever λ, μ. Also,

$$\Pi_0 = \left(1 + \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\right)^{-1} = \exp\{-\lambda/\mu).$$

We evaluate Π_k, using Eqn. (9), to get

$$\Pi_k = \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k \exp\{-\lambda/\mu\},\ k = 0, 1, 2, \ldots.$$

---x---