MICT-5101: Probability and Stochastic Process[1]

Dr. Md. Rezaul Karim
PhD(KULeuven & UHasselt), MS(Biostatistics), MS(Statistics)
Professor, Department of Statistics and Data Science
Jahangirnagar University (JU), Savar, Dhaka - 1342, Bangladesh

MS-2024

[1]These course slides should not be reproduced nor used by others (without permission).

Lecture Outline I

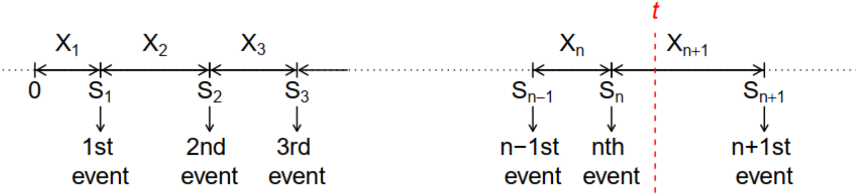
Chapter 7. Renewal Process and Branching Process

Renewal Process

We have seen that a Poisson process is a counting process for which the times between successive events are independent and identically distributed exponential random variables. One possible generalization is to consider a counting process for which the times between successive events are independent and identically distributed with an arbitrary distribution. Such a counting process is called a *renewal process*.

Definition 1.1

Let $\{N(t), t \geq 0\}$ be a counting process and let X_n denote the time between the $(n-1)$st and the nth event of this process, $n \geq 1$. If the sequence of nonnegative random variables $\{X_1, X_2, X_3 \ldots\}$ is independent and identically distributed, then the counting process $\{N(t), t \geq 0\}$ is said to be a renewal process. It was introduced as a generalization of Poisson process. It was considered as a process in continuous time.



Let $X_1, X_2, \ldots$ be i.i.d random variables with $\mathbb{E}[X_i] < \infty$, and $\mathrm{P}(X_i > 0) = 1$. Let

$$S_0 = 0, \quad S_n = X_1 + \ldots + X_n, \; n \geq 1.$$

Define

$$N(t) = \max\{n : S_n \leq t\}.$$

Then $\{N(t), t \geq 0\}$ is called a *renewal process*.

- Events are called "*renewals*". The interarrival times between events $X_1, X_2, \ldots$ are also called "renewals"
- A more general definition allows the first renewal X_1 to be of a different distribution, called a **delayed renewal process**.

Renewal Process

Example:

Suppose, we have an infinite number of electric light bulbs whose life times are iid random variables. Suppose we use a single light bulb at a time and we immediately replace it by a new one if it fail. Let $N(t)$ represents the number of light bulbs that have failed by time t, then $\{N(t); t \geq 0\}$ is a Renewal Process.

Renewal process used inn a variety of situations such as

- demography
- manpower studies
- reliability
- replacement and maintenance
- inventory control
- queuing process
- simulation and so on.

Classification of Renewal Process

1. Delayed/General/Modified renewal process
   A renewal process $\{N(t); t \geq 0\}$ is said to be delayed/general/modified renewal process if the first arrival time X_1 has a different distribution than the other $X_2, X_3, \ldots$.
2. Equilibrium renewal process
   A renewal process $\{N(t); t \geq 0\}$ is said to be equilibrium renewal process if the distribution of the 1st interarrival tie is specified which is

$$F_e(t) = \int_0^t \frac{1 - F(x)}{\mu} dx$$

where $\mu =$ mean of F. and it different from the iid distributed other.

Difference between Renewal Process and Poisson Process

(i) Interarrival times of the renewal process $\{N(t); g \geq 0\}$ are iid with arbitrary distribution.

(ii) The mean function is

$$m(t) = \sum_{n=1}^{\infty} F_n(t)$$

(iii) Renewal process is the generalization of Poisson process.

(iv) mean function may be linear.

(i) Interarriavals time have independently identical exponentially distributed.

(ii) The mean function is

$$m(t) = \lambda t$$

(iii) Renewal process is a particular case of renewal process.

(iv) Mean function always linear.

General Renewal Process: Distribution of $N(t)$

Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables representing the inter-arrival times in a renewal process. Define the total time of the n-th renewal by:

$$S_n = X_1 + X_2 + \cdots + X_n$$

The number of renewals $N(t)$ by time t is:

$$N(t) = \max\{n : S_n \le t\}$$

The probability that exactly n renewals occur by time t is:

$$P(N(t) = n) = P(S_n \le t \text{ and } S_{n+1} > t)$$

This can be written as the difference of cumulative distribution functions (CDFs) of S_n:

$$P(N(t) = n) = P(S_n \le t) - P(S_{n+1} \le t) = F_n(t) - F_{n+1}(t)$$

Example 7.1 Suppose that $P\{X_n = i\} = p(1-p)^{i-1}, i \geqslant 1$. That is, suppose that the interarrival distribution is geometric. Now $S_1 = X_1$ may be interpreted as the number of trials necessary to get a single success when each trial is independent and has a probability p of being a success. Similarly, S_n may be interpreted as the number of trials necessary to attain n successes, and hence has the negative binomial distribution

$$P\{S_n = k\} = \begin{cases} \binom{k-1}{n-1} p^n (1-p)^{k-n}, & k \geqslant n \\ 0, & k < n \end{cases}$$

Thus, from Equation (7.3) we have that

$$P\{N(t) = n\} = \sum_{k=n}^{[t]} \binom{k-1}{n-1} p^n (1-p)^{k-n} - \sum_{k=n+1}^{[t]} \binom{k-1}{n} p^{n+1} (1-p)^{k-n-1}$$

Equivalently, since an event independently occurs with probability p at each of the times 1, 2, ...

$$P\{N(t) = n\} = \binom{[t]}{n} p^n (1-p)^{[t]-n} \qquad \blacksquare$$

Another expression for $P(N(t) = n)$ can be obtained by conditioning on S_n. This yields

$$P\left(N(t) = n\right) = \int_0^\infty P\left(N(t) = n | S_n = y\right) f_{S_n}(y)dy$$

Now, if the nth event occurred at time $y > t$, then there would have been less than n events by time t. On the other hand, if it occurred at a time $y \leqslant t$, then there would be exactly n events by time t if the next interarrival exceeds $t - y$. Consequently,

$$\begin{aligned} P\left(N(t) = n\right) &= \int_0^t P\left(X_{n+1} > t - y | S_n = y\right) f_{S_n}(y)dy \\ &= \int_0^t \bar{F}(t-y) f_{S_n}(y)dy \end{aligned}$$

where $\bar{F} = 1 - F$.

Example 7.2 If $F(x) = 1 - e^{\lambda x}$ then S_n, being the sum of n independent exponentials with rate λ, will have a gamma (n, λ) distribution. Consequently, the preceding identity gives

$$
\begin{aligned}
P\left(N(t) = n\right) &= \int_0^t e^{-\lambda(t-y)} \frac{\lambda e^{-\lambda y} \left(\lambda y\right)^{n-1}}{(n-1)!} dy \\
&= \frac{\lambda^n e^{-\lambda t}}{(n-1)!} \int_0^t y^{n-1} dy \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!}
\end{aligned}
$$

■

Finding the Renewal Function

The mean value of a renewal process $\{N(t); t \geq 0\}$ is known as the renewal function. Let $P_n(t) = P(N(t) = n)$ is a probability mass function of $\{N(t); t \geq 0\}$. Then the expression of the renewal function is

$$
\begin{aligned}
m(t) &= E[N(t)] \\
&= \sum_{n=1}^{\infty} nP_n(t) \\
&= P_1(t) + 2P_2(t) + 3P_3(t) + 4P_4(t) + \cdots \\
&= P_1(t) + P_2(t) + P_3(t) + P_4(t) + \cdots \\
&\qquad\qquad + P_2(t) + P_3(t) + P_4(t) + \cdots \\
&\qquad\qquad\qquad\qquad + P_3(t) + P_4(t) + \cdots \\
&= \Pr(N(t) \geq 1) + \Pr(N(t) \geq 2) + \Pr(N(t) \geq 3) + \cdots \\
&= \sum_{n=1}^{\infty} \Pr(N(t) \geq n) = \sum_{n=1}^{\infty} \Pr(S_n \leq n) = \sum_{n=1}^{\infty} F_n(t)
\end{aligned}
$$

As $t \to \infty$, if the inter-arrival times have mean μ, the renewal process exhibits the following asymptotic behavior:

$$\lim_{t\to\infty} \frac{m(t)}{t} = \frac{1}{\mu}$$

Introduction to Branching Processes

- A **branching process** is a stochastic model that describes the evolution of populations where each individual in the population can produce offspring.
- Individuals reproduce and generate offspring, and the population size evolves over time.
- The key feature of a branching process is its probabilistic nature.
- The **Galton-Watson branching process** is a common model for such processes.

Key Features

- **Generations**: The population is divided into generations.
- **Offspring Distribution**: The number of offspring each individual has is determined by a random variable with a specific distribution.
- **Extinction**: The process may eventually result in the extinction of the population (i.e., no individuals are left).

Mathematical Model of a Branching Process

Consider a **Galton-Watson branching process**:

- Let Z_n be the population size at generation n.
- The population starts with Z_0 individuals.
- Each individual produces offspring independently according to some probability distribution p_k, where p_k is the probability that an individual produces k offspring.

The size of the population in the next generation is given by:

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_i$$

where X_i is the number of offspring produced by the i-th individual, and X_i are i.i.d. random variables.

Mean of Branching Process

Consider a Galton-Watson process where Z_n represents the population size at generation n.

The population at generation n is the sum of offspring from generation $n-1$:

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_i$$

where X_i is the number of offspring of the i-th individual, and X_i are i.i.d. random variables with mean $\mathbb{E}[X_i] = \mu$.

Using the linearity of expectation:

$$\mathbb{E}[Z_n] = \mathbb{E}\left[\sum_{i=1}^{Z_{n-1}} X_i\right] = \mathbb{E}[Z_{n-1}] \cdot \mathbb{E}[X_1] = \mu \cdot \mathbb{E}[Z_{n-1}]$$

Thus, iterating this recurrence:

$$\mathbb{E}[Z_n] = \mu^n \cdot Z_0$$

where Z_0 is the initial population size.

Variance of Branching Process

To calculate the variance $\mathrm{Var}(Z_n)$, we apply the law of total variance:

$$\mathrm{Var}(Z_n) = \mathbb{E}[\mathrm{Var}(Z_n \mid Z_{n-1})] + \mathrm{Var}(\mathbb{E}[Z_n \mid Z_{n-1}])$$

First, the conditional variance $\mathrm{Var}(Z_n \mid Z_{n-1})$ is:

$$\mathrm{Var}(Z_n \mid Z_{n-1}) = Z_{n-1} \cdot \mathrm{Var}(X_1)$$

where $\mathrm{Var}(X_1) = \sigma^2$, the variance in the number of offspring per individual. The conditional expectation $\mathbb{E}[Z_n \mid Z_{n-1}]$ is:

$$\mathbb{E}[Z_n \mid Z_{n-1}] = Z_{n-1} \cdot \mu$$

Therefore, the total variance is:

$$\mathrm{Var}(Z_n) = \sigma^2 \cdot \mathbb{E}[Z_{n-1}] + \mu^2 \cdot \mathrm{Var}(Z_{n-1})$$

Iterating this recurrence:

$$\mathrm{Var}(Z_n) = Z_0 \cdot \sigma^2 \cdot \frac{\mu^n - 1}{\mu - 1}$$

Types of Branching Processes

- **Supercritical Branching Process** ($\mu > 1$): The expected number of offspring per individual is greater than 1, and the population has a chance of growing indefinitely.
- **Critical Branching Process** ($\mu = 1$): The expected number of offspring per individual is equal to 1, and the population is expected to stay roughly constant.
- **Subcritical Branching Process** ($\mu < 1$): The expected number of offspring per individual is less than 1, and the population is expected to die out.

Applications of Branching Processes

- **Epidemiology**: Used to model the spread of infectious diseases in populations.
- **Population Biology**: Modeling the growth or extinction of species over time.
- **Physics**: Describing reactions and particle interactions in nuclear and particle physics.
- **Computer Science**: Studying recursive algorithms and data structures like search trees.