

Chapter 1

Preliminaries

1.1 Introduction

A fundamental concept in probability theory is random experiment. A random experiment is an experiment whose outcome cannot be determined in advance. The set of all possible outcomes of a random experiment is called the sample space of the experiment and is denoted by S .

An event E is a subset of sample space S . An event E is occurred if the outcome of the random experiment is an element of E . A random variable X is a real function defined on S .

A stochastic process $\{X(t), t \in T\}$ is a family of random variables so that for each $t \in T$, $X(t)$ is a random variable, where T is called the index set. We may interpret t as time, and then $X(t)$ is called the state of the stochastic process at time t . If the index set T is a countable set, the process $\{X(t), t \in T\}$ is called a discrete time stochastic process; if T is a continuum such, an interval for example, the process $\{X(t), t \in T\}$ is called a continuous time stochastic process.

Now, we introduce two concepts.

Definition 1.1.1. A stochastic process $\{X(t), t \in T\}$ is said to have independent increments if for all $t_0 < t_1 < \dots < t_n$ and $t_i \in T, i = 0, 1, \dots, n$, the random variables

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$$

are independent.

According to Definition 1.1.1, a stochastic process has independent increments if the changes of the process over nonoverlapping time intervals are independent.

Definition 1.1.2. A stochastic process $\{X(t), t \in T\}$ is said to have stationary increments if for all $t, t+s \in T$, the distribution of $X(t+s) - X(t)$ is the same for all t .

Thus, a stochastic process has stationary increments if the distribution of the change of the process between two times depends only on their time difference.

Chapter 1 introduces some preliminaries. In Section 1.2, we shall study the Poisson process that is a simple counting process. Section 1.3 will introduce the renewal process that is a more general counting process. In Section 1.4, we shall consider the stochastic order and discuss some classes of lifetime distribution. A new class of lifetime distribution, namely earlier preventive repair better (worse) than later in expectation, is introduced. Section 1.5 briefly studies the concept of martingale, then introduces martingale convergence theorem and the Doob, Riesz and Krickeberg composition theorems. In Section 1.6, we shall study the rate of occurrence of failures that is an important concept in reliability.

1.2 The Poisson Process

First of all, we define the counting process.

Definition 1.2.1. A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ is the total number of events occurred by time t .

Properties of the counting process.

- (1) $N(t) \geq 0$.
- (2) $N(t)$ is integer valued.
- (3) If $s < t$, then $N(s) \leq N(t)$.
- (4) For $s < t$, $N(t) - N(s)$ represents the number of events occurred in $(s, t]$.

It follows from Definition 1.1.1 that a counting process has independent increments if the numbers of events occurred in disjoint time intervals are independent. On the other hand, a counting process has a stationary increments if the distribution of the number of events occurred in a time interval depends only on the length of interval.

Now, we are available to define the Poisson process.

Definition 1.2.2. A counting process $\{N(t), t \geq 0\}$ is called a Poisson process with rate $\lambda > 0$, if

- (1) $N(0) = 0$.
- (2) The counting process has independent increments.
- (3) For all $s, t \geq 0$,

$$P\{N(t+s) - N(s) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad n = 0, 1, \dots \quad (1.2.1)$$

Note that (1.2.1) means that the number of events occurred in an time interval of length t is a Poisson random variable with mean λt .

Definition 1.2.3 gives an alternative definition of a Poisson process.

Definition 1.2.3. A counting process $\{N(t), t \geq 0\}$ is called a Poisson process with rate $\lambda > 0$, if

- (1) $N(0) = 0$.
- (2) The counting process has independent and stationary increments.
- (3) $P\{N(h) = 1\} = \lambda h + o(h)$.
- (4) $P\{N(h) \geq 2\} = o(h)$.

Theorem 1.2.4. Definitions 1.2.2 and 1.2.3 are equivalent.

Proof.

It is trivial to show that Definition 1.2.2 implies Definition 1.2.3. To show that Definition 1.2.3 implies Definition 1.2.2, let

$$p_n(t) = P\{N(t) = n\}.$$

By classical probability analysis, $p_n(t)$ will satisfy the following differential equations

$$p'_0(t) = -\lambda p_0(t), \quad (1.2.2)$$

$$p'_n(t) = -\lambda p_n(t) + \lambda p_{n-1}(t), \quad n = 1, 2, \dots \quad (1.2.3)$$

with initial condition

$$p_n(0) = \begin{cases} 1 & n = 0, \\ 0 & n \neq 0. \end{cases} \quad (1.2.4)$$

In fact, for $n > 0$,

$$\begin{aligned} p_n(t+h) &= P\{N(t+h) = n\} \\ &= P\{N(t) = n, N(t+h) - N(t) = 0\} \\ &\quad + P\{N(t) = n-1, N(t+h) - N(t) = 1\} \\ &\quad + \sum_{k=2}^n P\{N(t) = n-k, N(t+h) - N(t) = k\} \\ &= p_n(t)p_0(h) + p_{n-1}(t)p_1(h) + o(h) \end{aligned} \quad (1.2.5)$$

$$= (1 - \lambda h)p_n(t) + \lambda h p_{n-1}(t) + o(h), \quad (1.2.6)$$

where (1.2.5) is due to conditions (2) and (4), while (1.2.6) is in virtue of conditions (3) and (4) in Definition 1.2.3 respectively, so that

$$p_0(h) = 1 - P\{N(h) = 1\} - P\{N(h) \geq 2\} = 1 - \lambda h + o(h).$$

Therefore

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

Letting $h \rightarrow 0$ yields (1.2.3). The proof of (1.2.2) is similar. Now we can solve the differential equations with initial conditions. In fact, from (1.2.2) and (1.2.4) it is easy to see that

$$p_0(t) = e^{-\lambda t}.$$

Then assume that

$$p_{n-1}(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}.$$

From (1.2.3) with the help of (1.2.4), we have

$$p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}. \quad (1.2.7)$$

Thus by induction, (1.2.7) holds for all integers $n = 0, 1, \dots$. This proves that Definition 1.2.3 implies Definition 1.2.2. Hence, the proof of Theorem 1.2.4 is completed.

Definition 1.2.5. A continuous random variable X is said to have an exponential distribution $Exp(\lambda)$, if its density is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0, \\ 0 & \text{elsewhere.} \end{cases} \quad (1.2.8)$$

An important property of exponential distribution is the memoryless property. To explain this, consider

$$\begin{aligned} P\{X > s+t \mid X > s\} &= \frac{P\{X > s+t, X > s\}}{P\{X > s\}} = \frac{P\{X > s+t\}}{P\{X > s\}} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P\{X > t\}. \end{aligned} \quad (1.2.9)$$

This means that the conditional probability that a system survives for at least $s+t$ hours, given that it has survived for s hours, is the same as the unconditional probability that it survives for at least t hours. In other words, the system cannot remember how long it has survived. It can be

shown that a continuous distribution is memoryless if and only if it is an exponential distribution.

A discrete random variable X is said to have a geometric distribution $G(p)$ with parameter p , if the probability mass function is given by

$$p(x) = P\{X = x\} = pq^{x-1} \quad x = 1, 2, \dots \quad (1.2.10)$$

with $q = 1 - p$. Then $E\{X\} = 1/p$. An integer valued random variable X is memoryless if

$$P\{X > m + n \mid X > n\} = P\{X > m\} \quad \text{for } m, n = 0, 1, \dots \quad (1.2.11)$$

It can be shown that an integer valued distribution is memoryless if and only if it is a geometric distribution.

Given a Poisson process with rate λ , let X_1 be the time of the first event. In general, for $n \geq 1$, let X_n be the interarrival time between the $(n - 1)$ th and the n th events. Then we have the following theorem.

Theorem 1.2.6. Given a Poisson process with rate λ , the interarrival times $X_n, n = 1, 2, \dots$, are independent and identically distributed (i.i.d.) random variables each having an exponential distribution $Exp(\lambda)$.

Proof.

It is clear that

$$P\{X_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}.$$

Then X_1 has an exponential distribution $Exp(\lambda)$. Now consider

$$\begin{aligned} P\{X_1 > s, X_2 > t\} &= \int_0^\infty P\{X_1 > s, X_2 > t \mid X_1 = x\} \lambda e^{-\lambda x} dx \\ &= \int_s^\infty P\{\text{no event in } (x, t + x] \mid X_1 = x\} \lambda e^{-\lambda x} dx \\ &= \int_s^\infty P\{\text{no event in } (x, t + x)\} \lambda e^{-\lambda x} dx \quad (1.2.12) \end{aligned}$$

$$\begin{aligned} &= e^{-\lambda t} \int_s^\infty \lambda e^{-\lambda x} dx \quad (1.2.13) \\ &= e^{-\lambda s} e^{-\lambda t}, \end{aligned}$$

where (1.2.12) is due to independent increments, and (1.2.13) is in virtue to (1.2.1). Therefore, X_1 and X_2 are independent each having $Exp(\lambda)$ distribution. Then Theorem 1.2.4 follows by induction.

Now define $S_0 = 0$ and

$$S_n = \sum_{i=1}^n X_i,$$

S_n is the arrival time of the n th event. Clearly, we have

$$N(t) \geq n \Leftrightarrow S_n \leq t. \quad (1.2.14)$$

Because $X_i, i = 1, 2, \dots, n$, are i.i.d. random variables each having exponential distribution $Exp(\lambda)$. Then (1.2.14) yields that

$$P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!} e^{-\lambda t}. \quad (1.2.15)$$

Consequently, by differentiating with respect to t , S_n will have a gamma distribution $\Gamma(n, \lambda)$ with density function

$$f_n(x) = \begin{cases} \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} & x > 0, \\ 0 & \text{elsewhere.} \end{cases} \quad (1.2.16)$$

Assume that N is a Poisson random variable with parameter λ , and $\{X_i, i = 1, 2, \dots\}$ are i.i.d. random variables each having a distribution F . Assume further $\{X_i, i = 1, 2, \dots\}$ are independent of N . Then random variable

$$X = X_1 + \dots + X_N$$

is called a compound Poisson random variable with Poisson parameter λ and component distribution F .

Definition 1.2.7. Assume that $\{N(t), t \geq 0\}$ is a Poisson process, and $\{X_i, i = 1, 2, \dots\}$ are i.i.d. random variables and are independent of process $\{N(t), t \geq 0\}$. Let

$$X(t) = X_1 + \dots + X_{N(t)}.$$

Then stochastic process $\{X(t), t \geq 0\}$ is called a compound Poisson process.

Clearly, if $\{X(t), t \geq 0\}$ is a compound Poisson process, then $X(t)$ is a compound Poisson random variable with Poisson parameter λt .

In Definition 1.2.2 or 1.2.3, the rate λ of a Poisson process is a constant, hence the Poisson process is called a homogeneous Poisson process. Now, we shall consider the case that the rate is a function of time t . This is a nonhomogeneous Poisson process.

Definition 1.2.8. A counting process $\{N(t), t \geq 0\}$ is called a nonhomogeneous Poisson process with intensity function $\lambda(t), t \geq 0$, if

- (1) $N(0) = 0$.
- (2) The counting process has independent increments.
- (3) $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$.
- (4) $P\{N(t+h) - N(t) \geq 2\} = o(h)$.

Then we have the following theorem.

Theorem 1.2.9. Given a nonhomogeneous Poisson process $\{N(t), t \geq 0\}$ with intensity function $\lambda(t), t \geq 0$, then

$$P\{N(t+s) - N(s) = n\} = \frac{\left(\int_s^{t+s} \lambda(x) dx\right)^n}{n!} \exp\left\{-\int_s^{t+s} \lambda(x) dx\right\}$$

$$n = 0, 1, \dots \quad (1.2.17)$$

The proof is similar to that of Theorem 1.2.4, see Ross (1996) for details.

Now, let X be the lifetime until the first failure occurs. Then from (1.2.17), we have

$$\bar{F}(t) = P\{X > t\} = P\{N(t) = 0\} = \exp\left\{-\int_0^t \lambda(x) dx\right\}, \quad (1.2.18)$$

and $\bar{F}(t)$ is called the survival function. Thus the distribution F and the density f of X are respectively given by

$$F(x) = 1 - \exp\left\{-\int_0^x \lambda(u) du\right\}, \quad (1.2.19)$$

and

$$\lambda(x) = \frac{f(x)}{\bar{F}(x)}. \quad (1.2.20)$$

In other words, the intensity function $\lambda(t)$ and the distribution F are uniquely determined each other.

In practice, two particular nonhomogeneous Poisson processes are important.

(1) The Cox-Lewis model

The intensity function is

$$\lambda(t) = \exp(\alpha_0 + \alpha_1 t) \quad -\infty < \alpha_0, \alpha_1 < \infty, \quad t > 0. \quad (1.2.21)$$

(2) The Weibull process model

The intensity function is

$$\lambda(t) = \alpha\theta t^{\theta-1} \quad \alpha, \theta > 0, \quad t > 0. \quad (1.2.22)$$

1.3 The Renewal Process

Now, let $\{X_n, n = 1, 2, \dots\}$ be a sequence of nonnegative i.i.d. random variables each having a distribution F with $F(0) = P\{X_n = 0\} < 1$. Denote

$$\mu = E(X_n) = \int_0^{\infty} x dF(x).$$

Obviously $0 < \mu \leq \infty$. If we interpret X_n as the interarrival time between the $(n-1)$ th and n th events (or renewals), then as in Poisson process, we can define the arrival time of the n th event by

$$S_n = \sum_{i=1}^n X_i \quad (1.3.1)$$

with $S_0 = 0$. Thus, the number of events occurred by time t is then given by

$$N(t) = \sup\{n : S_n \leq t\}. \quad (1.3.2)$$

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

Definition 1.3.1. The counting process $\{N(t), t \geq 0\}$ is called a renewal process.

If the common distribution F is an exponential distribution, then the renewal process becomes a Poisson process. Therefore, renewal process is a generalization of Poisson process. Now define

$$M(t) = E[N(t)],$$

$M(t)$ is called the renewal function, it is the expected number of events occurred by time t .

As in Poisson process, we have

$$N(t) \geq n \iff S_n \leq t. \quad (1.3.3)$$

Then

$$\begin{aligned} P\{N(t) = n\} &= P\{N(t) \geq n\} - P\{N(t) \geq n + 1\} \\ &= P\{S_n \leq t\} - P\{S_{n+1} \leq t\} = F_n(t) - F_{n+1}(t), \end{aligned}$$

where F_n is the distribution function of S_n . Thus we have the following result.

Theorem 1.3.2.

$$M(t) = \sum_{n=1}^{\infty} F_n(t). \quad (1.3.4)$$

Proof.

$$\begin{aligned} M(t) &= E[N(t)] = \sum_{n=0}^{\infty} nP\{N(t) = n\} \\ &= \sum_{n=1}^{\infty} n\{F_n(t) - F_{n+1}(t)\} = \sum_{n=1}^{\infty} F_n(t). \end{aligned}$$

This completes the proof.

Theorem 1.3.3.

$$M(t) < \infty \quad \text{for all } 0 \leq t < \infty. \quad (1.3.5)$$

Proof.

Since $P\{X_n = 0\} < 1$, there exists an $\alpha > 0$ such that $p = P\{X_n \geq \alpha\} > 0$. Now define

$$\tilde{X}_n = \begin{cases} 0 & \text{if } X_n < \alpha, \\ \alpha & \text{if } X_n \geq \alpha. \end{cases} \quad (1.3.6)$$

Let

$$\tilde{S}_n = \tilde{X}_1 + \cdots + \tilde{X}_n, \quad \tilde{N}(t) = \sup\{n : \tilde{S}_n \leq t\}.$$

It is clear that $\{\tilde{N}(t), t \geq 0\}$ forms a renewal process whose events can only take place at times $t = k\alpha, k = 0, 1, \dots$. The number of events at

these times $t = k\alpha$ are i.i.d. random variables each having a geometric distribution $G(p)$. Thus

$$\begin{aligned} E[\tilde{N}(t)] &= \sum_{k=0}^{\lfloor t/\alpha \rfloor} E[\text{number of events occurred at } k\alpha] \\ &= \frac{\lfloor t/\alpha \rfloor + 1}{P(X_n \geq \alpha)} < \infty, \end{aligned}$$

where $\lfloor x \rfloor$ is the largest integer no more than x . As $\tilde{X}_n \leq X_n$, then we have $N(t) \leq \tilde{N}(t)$, and (1.3.5) follows.

Note that the above proof also shows that $E[(N(t))^r] < \infty$ for all $t \geq 0, r \geq 0$.

By conditional on the value of X_1 , we have

$$M(t) = E[N(t)] = \int_0^\infty E[N(t) \mid X_1 = x] dF(x). \tag{1.3.7}$$

Because

$$N(t) \mid X_1 = x = \begin{cases} 0 & x > t, \\ 1 + N_1(t - x) & x \leq t, \end{cases} \tag{1.3.8}$$

where

$$N_1(t) = \sup\{n : \sum_{i=2}^{n+1} X_i \leq t\}.$$

Then by substituting (1.3.8) to (1.3.7), it follows that

$$M(t) = \int_0^t \{1 + E[N(t - x)]\} dF(x),$$

since $N_1(t)$ and $N(t)$ have the same distribution. Thus we have

$$M(t) = F(t) + \int_0^t M(t - x) dF(x). \tag{1.3.9}$$

Equation (1.3.9) is called the renewal equation, it is an integral equation satisfied by renewal function $M(t)$.

Then assume that $\{X_n, n = 1, 2, \dots\}$ are continuous with common density function f . Let the density of S_n be f_n . Thus, Theorem 1.3.2 gives

$$M'(t) = \sum_{n=1}^\infty f_n(t). \tag{1.3.10}$$

Denote the Laplace transform of $M(t)$ by

$$M^*(s) = \int_0^\infty e^{-st} M(t) dt,$$

and the Laplace transform of $f(t)$ by

$$f^*(s) = \int_0^\infty e^{-st} f(t) dt.$$

Because $S_n = \sum_{i=1}^n X_i$ is the sum of X_1, \dots, X_n , the density of S_n is the n -fold convolution of f . Then the Laplace transform of $f_n(t)$ is given by

$$f_n^*(s) = [f^*(s)]^n.$$

Now, let the Laplace transform of $M'(t)$ be $M'^*(s)$. By taking Laplace transforms on both sides of (1.3.10), we have

$$M'^*(s) = \sum_{n=1}^\infty f_n^*(s) = \sum_{n=1}^\infty [f^*(s)]^n = \frac{f^*(s)}{1 - f^*(s)}.$$

However, as

$$M'^*(s) = sM^*(s) - M(0) = sM^*(s).$$

It follows from the above two equalities that

$$M^*(s) = \frac{f^*(s)}{s[1 - f^*(s)]}. \tag{1.3.11}$$

Thus, to determine renewal function $M(t)$, we can either use Theorem 1.3.2, or solve renewal equation (1.3.9). Alternatively, we can determine the Laplace transform $M^*(s)$ by (1.3.11) and then obtain $M(t)$ by inverting.

In many practical situations, an approximate formula for $M(t)$ is convenience. To derive an approximate formula, we need to introduce the concept of lattice.

A random variable X is lattice if it only takes on integral multiples of some $d \geq 0$. In other words, X is lattice, if there exists $d \geq 0$ such that

$$\sum_{n=-\infty}^\infty P\{X = nd\} = 1.$$

The largest d having this property is called the period of X . If X is lattice and F is the distribution of X , then we also say F is lattice.

Theorem 1.3.4. If F is not a lattice, $E[X] = \mu$ and $Var(X) = \sigma^2$, then

$$M(t) = \frac{t}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + o(1) \quad \text{as } t \rightarrow \infty. \quad (1.3.12)$$

Proof.

First of all, let the common density be f . Then the Laplace transform of f is given by

$$\begin{aligned} f^*(s) &= E[e^{-sX}] = E\left[1 - sX + \frac{1}{2}(sX)^2 - \dots\right] \\ &= 1 - s\mu + \frac{s^2}{2}(\sigma^2 + \mu^2) + O(s^3). \end{aligned} \quad (1.3.13)$$

By substitution, (1.3.11) becomes

$$\begin{aligned} M^*(s) &= \frac{1 - s\mu + s^2(\sigma^2 + \mu^2)/2 + O(s^3)}{s[s\mu - s^2(\sigma^2 + \mu^2)/2 + O(s^3)]} \\ &= \frac{1}{\mu s^2} [1 - s\mu + s^2(\sigma^2 + \mu^2)/2 + O(s^3)] \left[1 + \frac{\sigma^2 + \mu^2}{2\mu} s + O(s^2)\right] \\ &= \frac{1}{\mu s^2} + \frac{\sigma^2 - \mu^2}{2\mu^2 s} + O(1) \quad \text{as } s \rightarrow 0. \end{aligned} \quad (1.3.14)$$

Inverting (1.3.14) yields

$$M(t) = \frac{t}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + g(t) \quad \text{as } t \rightarrow \infty. \quad (1.3.15)$$

Let $g^*(s)$ be the Laplace transform of $g(t)$. Then from (1.3.14) and (1.3.15), by using the Tauberian theorem, we have

$$\begin{aligned} \lim_{t \rightarrow +\infty} g(t) &= \lim_{s \rightarrow 0} s g^*(s) \\ &= \lim_{s \rightarrow 0} s \left[M^*(s) - \frac{1}{\mu s^2} - \frac{\sigma^2 - \mu^2}{2\mu^2 s} \right] = 0. \end{aligned}$$

This completes the proof of Theorem 1.3.4.

Now, let $N(\infty) = \lim_{t \rightarrow \infty} N(t)$, then we have

Lemma 1.3.5.

$$P\{N(\infty) = \infty\} = 1. \quad (1.3.16)$$

Proof.

From (1.3.3), we have

$$\begin{aligned} P\{N(\infty) < \infty\} &= P\{X_n = \infty, \text{ for some } n\} \\ &= P\left\{\bigcup_{n=1}^{\infty} (X_n = \infty)\right\} \\ &\leq \sum_{n=1}^{\infty} P\{X_n = \infty\} = 0. \end{aligned}$$

Lemma 1.3.5 means that with probability 1, $N(t)$ tends to infinity as t tends to infinity. The following lemma gives the limiting rate of $N(t)$ as t tends to infinity.

Lemma 1.3.6. With probability 1,

$$\frac{N(t)}{t} \rightarrow \frac{1}{\mu} \quad \text{as } t \rightarrow \infty. \tag{1.3.17}$$

Proof.

Because $S_{N(t)} \leq t < S_{N(t)+1}$, then

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)}.$$

It follows from Lemma 1.3.5 that with probability 1, $N(t) \rightarrow \infty$ as $t \rightarrow \infty$. Then the strong law of large numbers shows that

$$\frac{S_{N(t)}}{N(t)} \rightarrow \mu \quad \text{as } t \rightarrow \infty.$$

On the other hand, we have

$$\begin{aligned} \frac{S_{N(t)+1}}{N(t)} &= \left\{ \frac{S_{N(t)+1}}{N(t)+1} \right\} \left\{ \frac{N(t)+1}{N(t)} \right\} \\ &\rightarrow \mu \quad \text{as } t \rightarrow \infty. \end{aligned}$$

As $t/N(t)$ is between two numbers, both of them converge to μ as $t \rightarrow \infty$, by principle of squeezing, so does $t/N(t)$. This completes the proof of Lemma 1.3.6.

Definition 1.3.7. Given a sequence of random variables $\{X_n, n = 1, 2, \dots\}$, an integer-valued random variable N is called a stopping time for $\{X_n, n = 1, 2, \dots\}$, if for all $n = 1, 2, \dots$, event $\{N = n\}$ is independent of X_{n+1}, X_{n+2}, \dots

Assume that N is a stopping time for stochastic process $\{X_n, n = 1, 2, \dots\}$. Then we can observe the process in sequential order and let N be the number observed before stopping. If $N = n$, then we shall stop after observing X_1, \dots, X_n and before observing X_{n+1}, X_{n+2}, \dots . Obviously, if N is a stopping time for $\{X_n, n = 1, 2, \dots\}$, then for all $n = 1, 2, \dots$ events $\{N \leq n\}$ and $\{N > n\}$ will be determined by X_1, \dots, X_n only. Now, we have the following theorem.

Theorem 1.3.8. (Wald's Equation)

If $\{X_n, n = 1, 2, \dots\}$ is a sequence of i.i.d. random variables each having a common expectation $E[X] < \infty$, and if N is a stopping time for $\{X_n, n = 1, 2, \dots\}$ with $E[N] < \infty$, then

$$E\left[\sum_{n=1}^N X_n\right] = E[N]E[X]. \quad (1.3.18)$$

Proof.

Let the indicator of event $\{N \geq n\}$ be

$$I_n = \begin{cases} 1 & \text{if } N \geq n, \\ 0 & \text{if } N < n. \end{cases} \quad (1.3.19)$$

Then

$$\begin{aligned} E\left[\sum_{n=1}^N X_n\right] &= E\left[\sum_{n=1}^{\infty} X_n I_n\right] = \sum_{n=1}^{\infty} E[X_n I_n] \\ &= E[X] \sum_{n=1}^{\infty} E[I_n] \end{aligned} \quad (1.3.20)$$

$$= E[X] \sum_{n=1}^{\infty} P\{N \geq n\} = E[N]E[X], \quad (1.3.21)$$

where (1.3.20) is because N is a stopping time, then event $\{N \geq n\}$ is determined by X_1, \dots, X_{n-1} and hence independent of X_n . This completes the proof.

Now, assume that $\{N(t), t \geq 0\}$ is a renewal process whose interarrival times are $\{X_n, n = 1, 2, \dots\}$. Then $N(t) + 1$ is a stopping time for $\{X_n, n = 1, 2, \dots\}$. This is because

$$\begin{aligned} N(t) + 1 = n &\Leftrightarrow N(t) = n - 1 \\ &\Leftrightarrow X_1 + \dots + X_{n-1} \leq t, \quad X_1 + \dots + X_n > t. \end{aligned}$$

Thus event $\{N(t) + 1 = n\}$ is determined by X_1, \dots, X_n . Consequently, $N(t) + 1$ is a stopping time. Therefore Theorem 1.3.8 yields the following

result.

Corollary 1.3.9. If $\mu < \infty$, then

$$E[S_{N(t)+1}] = \mu\{M(t) + 1\}. \tag{1.3.22}$$

The following theorems are important in renewal process.

Theorem 1.3.10. (The Elementary Renewal Theorem)

$$\frac{M(t)}{t} \rightarrow \frac{1}{\mu} \text{ as } t \rightarrow \infty, \tag{1.3.23}$$

where $1/\infty = 0$.

See Ross (1996) for the proof.

Theorem 1.3.11. Let μ and σ^2 be the mean and variance of the common distribution of the interarrival time in a renewal process $\{N(t), t \geq 0\}$. Then

$$P\left\{\frac{N(t) - t/\mu}{\sigma\sqrt{t/\mu^3}} < y\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx \text{ as } t \rightarrow \infty. \tag{1.3.24}$$

Proof.

Let $r_t = t/\mu + y\sigma\sqrt{t/\mu^3}$. Then by using (1.3.3)

$$\begin{aligned} P\left\{\frac{N(t) - t/\mu}{\sigma\sqrt{t/\mu^3}} < y\right\} &= P\{N(t) < r_t\} \\ &= P\{S_{r_t} > t\} = P\left\{\frac{S_{r_t} - r_t\mu}{\sigma\sqrt{r_t}} > \frac{t - r_t\mu}{\sigma\sqrt{r_t}}\right\} \\ &= P\left\{\frac{S_{r_t} - r_t\mu}{\sigma\sqrt{r_t}} > -y\left(1 + \frac{y\sigma}{\sqrt{t\mu}}\right)^{-1/2}\right\} \\ &\rightarrow \frac{1}{\sqrt{2\pi}} \int_{-y}^{\infty} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx \text{ as } t \rightarrow \infty, \end{aligned} \tag{1.3.25}$$

(1.3.25) is due to the central limit theorem so that $(S_{r_t} - r_t\mu)/\sigma\sqrt{r_t}$ converges to standard normal distribution $N(0, 1)$ as t tends to infinity, and because

$$-y\left(1 + \frac{y\sigma}{\sqrt{t\mu}}\right)^{-1/2} \rightarrow -y \text{ as } t \rightarrow \infty.$$

This completes the proof.

Theorem 1.3.11 states that if t is large, then $N(t)$ will approximately be normal with mean t/μ and variance $t\sigma^2/\mu^3$.

Theorem 1.3.12. (Blackwell’s Theorem)

Let F be the common distribution of interarrival times in a renewal process.

(1) If F is not lattice, then for all $a \geq 0$

$$M(t + a) - M(t) \rightarrow a/\mu \text{ as } t \rightarrow \infty. \tag{1.3.26}$$

(2) If F is lattice with period d , then

$$E[\text{number of renewals at } nd] \rightarrow d/\mu \text{ as } t \rightarrow \infty. \tag{1.3.27}$$

For the proof of Theorem 1.3.12, see e.g. Feller (1970).

Given a function h defined on $[0, \infty)$, for any $a > 0$, let $\overline{m}_n(a)$ and $\underline{m}_n(a)$ be the supremum and infimum of $h(t)$ on interval $[(n - 1)a, na]$ respectively. Then h is said to be directly Riemann integrable if for any $a > 0$, $\overline{m}_n(a)$ and $\underline{m}_n(a)$ are finite and

$$\lim_{a \rightarrow 0} \{a \sum_{n=1}^{\infty} \overline{m}_n(a)\} = \lim_{a \rightarrow 0} \{a \sum_{n=1}^{\infty} \underline{m}_n(a)\}.$$

A sufficient condition for a function h to be directly Riemann integrable is that

- (1) $h(t) \geq 0$ for all $t \geq 0$,
- (2) $h(t)$ is nonincreasing,
- (3) $\int_0^{\infty} h(t)dt < \infty$.

Then we have the following theorem.

Theorem 1.3.13. (The Key Renewal Theorem)

Let F be the common distribution of the interarrival times in a renewal process. If F is not lattice, and h is directly Riemann integrable, then

$$\lim_{t \rightarrow \infty} \int_0^t h(t - x)dM(x) = \frac{1}{\mu} \int_0^{\infty} h(t)dt, \tag{1.3.28}$$

where

$$M(x) = \sum_{n=1}^{\infty} F_n(x) \quad \text{and} \quad \mu = \int_0^{\infty} \bar{F}(t)dt.$$

For the proof of Theorem 1.3.13, see e.g. Feller (1970).

Given a renewal process $\{N(t), t \geq 0\}$ with interarrival times $\{X_n, n = 1, 2, \dots\}$, we can define the age at t as

$$A(t) = t - S_{N(t)}, \tag{1.3.29}$$

the residual life at t as

$$B(t) = S_{N(t)+1} - t, \tag{1.3.30}$$

and the total life at t as

$$X_{N(t)+1} = S_{N(t)+1} - S_{N(t)} = A(t) + B(t). \tag{1.3.31}$$

Then we have the following theorem.

Theorem 1.3.14. Let F be the common distribution of the interarrival times in a renewal process. If F is not lattice with mean $\mu < \infty$, then

$$(1) \lim_{t \rightarrow \infty} P\{A(t) \leq x\} = \lim_{t \rightarrow \infty} P\{B(t) \leq x\} = \frac{1}{\mu} \int_0^x \bar{F}(y) dy, \tag{1.3.32}$$

$$(2) \lim_{t \rightarrow \infty} P\{X_{N(t)+1} \leq x\} = \frac{1}{\mu} \int_0^x y dF(y). \tag{1.3.33}$$

See Ross (1996) for the proof of Theorem 1.3.14.

Now, consider a renewal process $\{N(t), t \geq 0\}$ with interarrival times $\{X_n, n = 1, 2, \dots\}$ having a common distribution F . Assume that each time a renewal occurs we shall receive a reward. Let R_n be the reward earned at the time of the n th renewal. Usually, R_n will depend on X_n , but we assume that the pairs $\{(X_n, R_n), n = 1, 2, \dots\}$ are i.i.d. random vectors. Then $\{(X_n, R_n), n = 1, 2, \dots\}$ is called a renewal reward process. Thus, the total reward earned by time t is given by

$$R(t) = \sum_{n=1}^{N(t)} R_n.$$

Let

$$E[R] = E[R_n], \quad E[X] = E[X_n].$$

Afterward, we have the following theorem.

Theorem 1.3.15. Assume that $E[R] < \infty$ and $E[X] < \infty$. Then with probability 1,

$$\frac{R(t)}{t} \rightarrow \frac{E[R]}{E[X]} \quad \text{as } t \rightarrow \infty, \tag{1.3.34}$$

and

$$\frac{E[R(t)]}{t} \rightarrow \frac{E[R]}{E[X]} \quad \text{as } t \rightarrow \infty. \quad (1.3.35)$$

See Ross (1996) for the proof of Theorem 1.3.15. This is an important result and has a lot of applications in practice.

Remarks

If we say a cycle is completed whenever a renewal occurs, then (1.3.35) states that the long-run average reward per unit time (or simply average reward) is just the expected reward earned during a cycle, divided by the expected time of a cycle.

The reward can be negative. In this case, we may define R_n as the cost incurred at the time of the n th renewal. Then (1.3.35) states that the long-run average cost per unit time (or simply average cost) is given by

$$\text{average cost} = \frac{\text{the expected cost incurred in a cycle}}{\text{the expected time of a cycle}}. \quad (1.3.36)$$

Although we have assumed that the reward is earned all at once at the end of the renewal cycle, Theorem 1.3.15 still holds when the reward is earned gradually during the renewal cycle.

1.4 Stochastic Order and Class of Lifetime Distributions

At first, we introduce the concept of stochastic order here.

Definition 1.4.1. A random variable X is said to be stochastically larger (less) than a random variable Y , if

$$P\{X > t\} \geq (\leq) P\{Y > t\} \quad \text{for all } t. \quad (1.4.1)$$

It is denoted by $X \geq_{st} (\leq_{st}) Y$. Let F and G be the distributions of X and Y respectively, then (1.4.1) is equivalent to

$$\bar{F}(t) \geq (\leq) \bar{G}(t) \quad \text{for all } t. \quad (1.4.2)$$

A sequence of random variables $\{X_n, n = 1, 2, \dots\}$ is stochastically increasing (decreasing) if

$$X_{n+1} \geq_{st} (\leq_{st}) X_n, \quad n = 1, 2, \dots$$

Lemma 1.4.2. If $X \geq_{st} Y$, then $E[X] \geq E[Y]$.

Proof.

Assume first that X and Y are nonnegative random variables. Then

$$E[X] = \int_0^{\infty} P\{X > t\} dt \geq \int_0^{\infty} P\{Y > t\} dt = E[Y]. \quad (1.4.3)$$

In general, we have

$$X = X^+ - X^-,$$

where

$$X^+ = \begin{cases} X & \text{if } X \geq 0, \\ 0 & \text{if } X < 0, \end{cases} \quad (1.4.4)$$

and

$$X^- = \begin{cases} 0 & \text{if } X \geq 0, \\ -X & \text{if } X < 0. \end{cases} \quad (1.4.5)$$

Similarly, $Y = Y^+ - Y^-$. It is easy to show that

$$X \geq_{st} Y \Rightarrow X^+ \geq_{st} Y^+, \quad X^- \leq_{st} Y^-. \quad (1.4.6)$$

Consequently, it follows from (1.4.3) and (1.4.6) that

$$E[X] = E[X^+] - E[X^-] \geq E[Y^+] - E[Y^-] = E[Y]. \quad (1.4.7)$$

The following theorem gives a sufficient and necessary condition so that X is stochastically larger than Y .

Theorem 1.4.3.

$$X \geq_{st} Y \iff E[f(X)] \geq E[f(Y)]$$

for any nondecreasing function f . (1.4.8)

Proof.

Suppose that $X \geq_{st} Y$ and f is a nondecreasing function. For any t , let

$$f^{-1}(t) = \inf\{x : f(x) > t\}.$$

Then

$$P\{f(X) > t\} = P\{X > f^{-1}(t)\} \geq P\{Y > f^{-1}(t)\} = P\{f(Y) > t\}.$$

Consequently, $F(X) \geq_{st} f(Y)$. Then, from Lemma 1.4.2, $E[f(X)] \geq E[f(Y)]$.

Now assume that $E[f(X)] \geq E[f(Y)]$ for any nondecreasing function f . For any t , define

$$f_t(x) = \begin{cases} 1 & x > t, \\ 0 & \text{elsewhere.} \end{cases}$$

Then

$$E[f_t(X)] = P\{X > t\}, \quad E[f_t(Y)] = P\{Y > t\}. \quad (1.4.9)$$

Because $E[f_t(X)] \geq E[f_t(Y)]$, (1.4.9) implies that $X \geq_{st} Y$.

Then assume that X is the lifetime of a system with distribution F . The residual life of the system at age t given that the system has survived up time t will be given by

$$X_t = X - t \mid X > t. \quad (1.4.10)$$

A very popular class of life distribution is studied on the basis of the monotonicity of X_t .

Definition 1.4.4. A life distribution F has an increasing failure rate (IFR) if X_t is decreasing in t , i.e.

$$X_s \geq_{st} X_t \quad \text{for } s < t. \quad (1.4.11)$$

Or for any $x \geq 0$, we have

$$P(X - s > x \mid X > s) \geq P(X - t > x \mid X > t) \quad \text{for } s < t. \quad (1.4.12)$$

It has a decreasing failure rate (DFR) if X_t is increasing in t , i.e.

$$X_s \leq_{st} X_t \quad \text{for } s < t. \quad (1.4.13)$$

Or for any $x \geq 0$, we have

$$P(X - s > x \mid X > s) \leq P(X - t > x \mid X > t) \quad \text{for } s < t. \quad (1.4.14)$$

If F has an IFR, it is denoted by $F \in \text{IFR}$, if F has a DFR, it is denoted by $F \in \text{DFR}$. Now define

$$\bar{F}_t(x) = P(X - t > x \mid X > t) = \frac{\bar{F}(t+x)}{\bar{F}(t)}.$$

Then we have the following result.

Theorem 1.4.5. $F \in \text{IFR}$ (DFR) if and only if $\bar{F}_t(x)$ is decreasing (increasing) in t .

The class of IFR (DFR) distributions is closely related to the concept of failure rate. Consider a continuous random variable X , let the distribution and density be F and f respectively. Then the failure rate function is defined by

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}. \tag{1.4.15}$$

Let X be the lifetime of a system. Given that the system has survived for time t , the conditional probability that it will not survive for an additional time dt is given by

$$\begin{aligned} P\{t < X < t + dt \mid X > t\} &= \frac{P\{t < X < t + dt, X > t\}}{P\{X > t\}} \\ &= \frac{P\{t < X < t + dt\}}{P\{X > t\}} = \frac{\int_t^{t+dt} f(u)du}{\bar{F}(t)} \\ &\doteq \frac{f(t)dt}{\bar{F}(t)} = \lambda(t)dt. \end{aligned} \tag{1.4.16}$$

Thus $\lambda(t)$ is the probability intensity that a t -year-old system will fail in $(t, t + dt)$.

Now, assume that X is the lifetime until the first failure in a nonhomogeneous Poisson process with intensity function $\lambda(t)$. Then it follows from (1.2.19), (1.2.20) and (1.4.15) that intensity function $\lambda(t)$ is the failure rate function of X .

In particular, if X is the lifetime before the first failure in a Poisson process with rate λ , then lifetime X will have an exponential distribution $Exp(\lambda)$ with density function given by (1.2.8). As a result, the failure rate is

$$\lambda(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda.$$

Therefore, the failure rate of an exponential distribution $Exp(\lambda)$ is a constant λ that is the rate of the Poisson process.

Let

$$d = \sup\{t : t \geq 0, \bar{F}(t) > 0\}.$$

It can be shown that if $F \in \text{IFR(DFR)}$, then F is absolutely continuous in $(0, d)$, i.e. F has a probability density in $(0, d)$. See Barlow and Proschan (1975) or Cheng (1999) for reference. Now we have the following equality.

$$\begin{aligned} \frac{d\bar{F}_t(x)}{dt} &= \frac{1}{\{\bar{F}(t)\}^2} \{-f(t+x)\bar{F}(t) + f(t)\bar{F}(t+x)\} \\ &= \frac{\bar{F}(t+x)}{\bar{F}(t)} \{\lambda(t) - \lambda(t+x)\}. \end{aligned}$$

Then we have the following theorem.

Theorem 1.4.6. A continuous life distribution F has an IFR (DFR) if and only if its failure rate function $\lambda(t)$ is increasing (decreasing).

From Theorem 1.4.6, we can see intuitively that X is IFR (DFR) means that the older the system is, the more (less) likely to fail in a small time dt is.

The following results are well known, see Barlow and Proschan (1975) for reference.

Theorem 1.4.7.

(1) A gamma distribution $\Gamma(\alpha, \beta)$ with density function

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0, \\ 0 & \text{elsewhere,} \end{cases} \quad (1.4.17)$$

is IFR if $\alpha \geq 1$ and DFR if $\alpha \leq 1$.

(2) A Weibull distribution $W(\alpha, \beta)$ with density function

$$f(x) = \begin{cases} \alpha\beta x^{\alpha-1} e^{-\beta x^\alpha} & x > 0, \\ 0 & \text{elsewhere,} \end{cases} \quad (1.4.18)$$

is IFR if $\alpha \geq 1$ and DFR if $\alpha \leq 1$.

In particular, an exponential distribution $Exp(\lambda)$ with density function (1.2.8) is IFR as well as DFR.

The set of all IFR (DFR) distributions forms a class of life distribution. Now, we consider another class of life distribution.

Definition 1.4.8. A life distribution F with finite mean μ is new better than used in expectation (NBUE) if

$$\bar{F}(t) \geq \frac{1}{\mu} \int_t^\infty \bar{F}(x) dx \quad \text{for } t \geq 0. \quad (1.4.19)$$

It is new worse than used in expectation (NWUE) if

$$\bar{F}(t) \leq \frac{1}{\mu} \int_t^\infty \bar{F}(x) dx \quad \text{for } t \geq 0. \quad (1.4.20)$$

They are denoted respectively by $F \in \text{NBUE}$ and $F \in \text{NWUE}$.

The set of all NBUE (NWUE) distributions forms a class of life distributions. The following lemma gives the relationship between IFR (DFR)

and NBUE (NWUE) classes.

Lemma 1.4.9. If a lifetime distribution $F \in \text{IFR}$ (DFR), then $F \in \text{NBUE}$ (NWUE).

See Barlow and Proschan (1975) for the proof.

Let the lifetime of a system be X . Given that $X > t$, then $X_t = X - t$ is the residual life. Thus from Definition 1.4.8, we can see that a life distribution F is NBUE (NWUE) if the expected residual life of a system which has survived for time t is no larger (less) than the expected lifetime of a new and identical system. In fact, if $F \in \text{NBUE}$ (NWUE), then from Definition 1.4.8, it yields that

$$\begin{aligned} E[X_t] &= \int_0^{\infty} x dF_t(x) = \frac{1}{\bar{F}(t)} \int_0^{\infty} x dF(t+x) \\ &= \frac{1}{\bar{F}(t)} \int_t^{\infty} \bar{F}(x) dx \leq (\geq) \mu. \end{aligned}$$

Besides IFR (DFR) and NBUE (NWUE) classes, many other classes of life distributions have been considered (see Barlow and Proschan (1975) and Cheng (1999) for reference). However, all the classes of life distributions studied so far are defined on the basis of the life distribution of a system itself, i.e. the ageing effect on the system itself. In practice, for improving the system reliability and implementing the system more economically, besides the failure repair, we may also adopt a preventive repair that is a repair during the operating time of the system. Then, it is interesting to study the effect of preventive repair on the distribution of the total operating time of the system. For this purpose, we shall introduce a new class of lifetime distribution by taking into account the effect of preventive repair.

Let $F(x)$ be the distribution of the operating time X of a system. Assume that a preventive repair is taken when the operating time of the system reaches T , and the system after preventive repair is as good as new. Let M be the number of preventive repairs before system failure. Then, $Y = M + 1$ will have a geometric distribution $G(p)$ with probability mass function given by (1.2.10), where

$$p = P\{X \leq T\} = F(T)$$

and $q = P\{X > t\} = 1 - p$. Moreover,

$$E[M] = E(Y) - 1 = \frac{1}{p} - 1 = \frac{q}{p}.$$

Now, the total operating time with preventive repair at T is given by

$$X_T = \sum_{i=1}^M T + X_{(M+1)} = MT + X_{(M+1)},$$

where $X_{(M+1)} < T$ is the $(M + 1)$ th operating time of the system. Thus, the expected total operating time with preventive repair at T is given by

$$\begin{aligned} E[X_T] &= E\{E[X_T | M]\} \\ &= E\{E[MT + X_{(M+1)} | M]\} \\ &= E\left\{MT + \frac{1}{F(T)} \int_0^T t dF(t)\right\} \\ &= \frac{1}{p} \left\{qT + \int_0^T t dF(t)\right\} \\ &= \frac{1}{F(T)} \int_0^T \{1 - F(t)\} dt. \end{aligned} \quad (1.4.21)$$

Then, Lam (2006b) introduced the following definition.

Definition 1.4.10. A life distribution F is called earlier preventive repair better than later in expectation (ERBLE) if the expected total operating time with preventive repair at t

$$\mu(t) = \frac{1}{F(t)} \int_0^t \{1 - F(x)\} dx \quad (1.4.22)$$

is nonincreasing in t , this is denoted by $F \in \text{ERBLE}$. It is called earlier preventive repair worse than later in expectation (ERWLE) if $\mu(t)$ is non-decreasing in t , and is denoted by $F \in \text{ERWLE}$.

The following results are also due to Lam (2006b).

Lemma 1.4.11. If a life distribution $F \in \text{IFR}$ (DFR), then $F \in \text{ERBLE}$ (ERWLE).

Proof.

If $F \in \text{IFR}$, F will be absolutely continuous. Let the density function be f , then the failure rate function is given by $\lambda(t) = f(t)/\{1 - F(t)\}$. Thus

it follows from (1.4.22) that

$$\begin{aligned} \mu'(t) &= \frac{1}{F(t)^2} \{F(t)[1 - F(t)] - f(t) \int_0^t [1 - F(x)]dx\} \\ &= \frac{f(t)}{F(t)^2} \left\{ \frac{F(t)}{\lambda(t)} - \int_0^t \frac{f(x)}{\lambda(x)} dx \right\} \\ &\leq \frac{f(t)}{\lambda(t)F(t)^2} \{F(t) - \int_0^t f(x)dx\} = 0. \end{aligned}$$

Thus, $F \in \text{ERBLE}$. For the case $F \in \text{DFR}$, the proof is similar.

Then, from Theorem 1.4.7, we have

Corollary 1.4.12.

(1) A gamma distribution $\Gamma(\alpha, \beta)$ with density function (1.4.17) is ERBLE if $\alpha \geq 1$ and ERWLE if $\alpha \leq 1$.

(2) A Weibull distribution $W(\alpha, \beta)$ with density function (1.4.18) is ERBLE if $\alpha \geq 1$ and ERWLE if $\alpha \leq 1$.

In particular, an exponential distribution $\text{Exp}(\lambda)$ with density function (1.2.8) is ERBLE as well as ERWLE.

To see the relation between the classes of ERBLE (ERWLE) and IFR (DFR). we can consider the following two examples.

Example 1.4.13. Let

$$F(x) = \begin{cases} 1 - e^{-x} & 0 \leq x < 1, \\ 1 - e^{-3x} & x \geq 1. \end{cases} \tag{1.4.23}$$

It is easy to check that for $s < 1 < t$ and $1 < s + x < t + x$, then

$$\begin{aligned} \bar{F}_s(x) &= \frac{\bar{F}(s+x)}{\bar{F}(s)} = e^{-(2s+3x)} \\ &< \bar{F}_t(x) = \frac{\bar{F}(t+x)}{\bar{F}(t)} = e^{-3x}. \end{aligned}$$

Thus $F \notin \text{IFR}$. However, for $0 \leq t < 1, \mu(t) = 1$, it is a constant. Moreover, for $t \geq 1$

$$\mu(t) = \frac{1}{1 - e^{-3t}} (1 - e^{-1} + \frac{1}{3}e^{-3} - \frac{1}{3}e^{-3t}),$$

it is decreasing. This implies that $F \in \text{ERBLE}$.

Example 1.4.14. Let

$$F(x) = \begin{cases} 1 - e^{-1} & 0 \leq x < 1, \\ 1 - e^{-x} & x \geq 1. \end{cases} \tag{1.4.24}$$

By a similar argument, it is also easy to show that $F \in \text{ERWLE}$, but $F \notin \text{DFR}$.

From Lemma 1.4.11 and two examples above, we can conclude that the ERBLE class (ERWLE) is greater than the IFR (DFR) class. The class of ERBLE (ERWLE) is a new class of life distributions by taking into account the effect of preventive repair. This new class of life distributions should have its own theoretical interest and practical application.

1.5 * Martingales

Consider a stochastic process $\{Z_n, n = 1, 2, \dots\}$, let

$$\mathcal{F}_n = \sigma\{Z_1, \dots, Z_n\}$$

be the σ -algebra generated by $\{Z_1, \dots, Z_n\}$. If we interpret Z_n as the system state at time n , then \mathcal{F}_n will represent the history of the system up to time n .

A stochastic process $\{X_n, n = 1, 2, \dots\}$ is called a martingale with respect to $\{\mathcal{F}_n, n = 1, 2, \dots\}$ if for all n

$$E[|X_n|] < \infty, \quad (1.5.1)$$

$$X_n \in \mathcal{F}_n \text{ i.e. } X_n \text{ is } \mathcal{F}_n \text{ measurable}, \quad (1.5.2)$$

$$E[X_{n+1} | \mathcal{F}_n] = X_n. \quad (1.5.3)$$

If $\{X_n, n = 1, 2, \dots\}$ is a martingale, then taking expectation on both sides of (1.5.3) yields

$$E[X_{n+1}] = E[X_n],$$

and hence

$$E[X_n] = E[X_1] \text{ for all } n. \quad (1.5.4)$$

A stochastic process $\{X_n, n = 1, 2, \dots\}$ is called a supermartingale (submartingale) with respect to $\{\mathcal{F}_n, n = 1, 2, \dots\}$ if for all n , conditions (1.5.1), (1.5.2) and

$$E[X_{n+1} | \mathcal{F}_n] \leq (\geq) X_n, \quad (1.5.5)$$

hold. If $\{X_n, n = 1, 2, \dots\}$ is a supermartingale (submartingale), then taking expectation of (1.5.5) yields

$$E[X_{n+1}] \leq (\geq) E[X_n],$$

and hence

$$E[X_n] \leq (\geq) E[X_1] \text{ for all } n. \tag{1.5.6}$$

Theorem 1.5.1. (Martingale Convergence Theorem)

Assume that $\{X_n, n = 1, 2, \dots\}$ is a supermartingale (submartingale) with

$$\sup_{1 \leq n < \infty} E[|X_n|] < \infty.$$

Then with probability 1, random variable

$$X = \lim_{n \rightarrow \infty} X_n$$

exists and $E[|X|] < \infty$. See e. g. Stout (1974) for the proof of Theorem 1.5.1.

Theorem 1.5.2. (The Doob Decomposition)

Assume that $\{X_n, n = 1, 2, \dots\}$ is a submartingale with respect to $\{\mathcal{F}_n, n = 1, 2, \dots\}$. Then $\{X_n, n = 1, 2, \dots\}$ has a decomposition such that

$$X_n = M_n - A_n, \tag{1.5.7}$$

where $M_n, n = 1, 2, \dots$ is a martingale, $\{A_n, n = 1, 2, \dots\}$ is an decreasing sequence of random variables such that $A_{n+1} \leq A_n$ with probability 1, $A_1 = 0$ and $A_n \in \mathcal{F}_{n-1}$. Moreover, such a decomposition is unique.

Proof.

Let $M_1 = X_1$ and $A_1 = 0$. Then define M_n, A_n for $n \geq 2$ as follows.

$$M_n = M_{n-1} + (X_n - E[X_n | \mathcal{F}_{n-1}]), \tag{1.5.8}$$

$$A_n = A_{n-1} + (X_{n-1} - E[X_n | \mathcal{F}_{n-1}]). \tag{1.5.9}$$

From (1.5.8) and (1.5.9), we have

$$M_n - A_n = \sum_{i=2}^n (X_i - X_{i-1}) + X_1 - A_1 = X_n,$$

and (1.5.7) follows. It is then easy to check $M_n, n = 1, 2, \dots$ and $\{A_n, n = 1, 2, \dots\}$ satisfy the requirements. To show such a decomposition is unique, suppose $X_n = M'_n - A'_n$ is another decomposition. Then

$$M_n - M'_n = A_n - A'_n.$$

Thus $M_1 = M'_1$, since $A_1 = A'_1 = 0$. Moreover, because $M_2 - M'_2 = A_2 - A'_2 \in \mathcal{F}_1$, therefore,

$$M_2 - M'_2 = E[M_2 - M'_2 | \mathcal{F}_1] = M_1 - M'_1 = 0.$$

Then by induction, we can show that $M_n = M'_n$ and hence $A_n = A'_n$ for all n . This completes the proof of Theorem 1.5.2.

Theorem 1.5.3. (The Riesz Decomposition)

Assume that $\{X_n, n = 1, 2, \dots\}$ is a submartingale with respect to $\{\mathcal{F}_n, n = 1, 2, \dots\}$. Then the following two conditions are equivalent.

$$(1) \quad \lim_{n \rightarrow \infty} E[X_n] < \infty, \tag{1.5.10}$$

$$(2) \quad \{X_n, n = 1, 2, \dots\} \text{ has a decomposition such that} \\ X_n = Y_n + Z_n, \tag{1.5.11}$$

where $\{Y_n, n = 1, 2, \dots\}$ is a martingale and $\{Z_n, n = 1, 2, \dots\}$ is a non-positive submartingale with $\lim_{n \rightarrow \infty} E[Z_n] = 0$. Moreover, such a decomposition is unique. If in addition, $\{X_n, n = 1, 2, \dots\}$ are nonnegative, then $\{Y_n, n = 1, 2, \dots\}$ are also nonnegative.

Proof.

(1) \Rightarrow (2) By Theorem 1.5.2, $\{X_n, n = 1, 2, \dots\}$ has a Doob decomposition $X_n = M_n - A_n$, then

$$E[X_n] = E[M_n] - E[A_n] = E[M_1] - E[A_n].$$

As A_n is nonpositive, then $E[|A_n|] = E[X_n] - E[M_1] < \infty$, and $A_\infty = \lim_{n \rightarrow \infty} A_n$ is integrable. Define

$$Y_n = M_n - E[A_\infty | \mathcal{F}_n], \tag{1.5.12}$$

$$Z_n = E[A_\infty | \mathcal{F}_n] - A_n. \tag{1.5.13}$$

Now $\{Y_n, n = 1, 2, \dots\}$ is a martingale, since $\{E[A_\infty | \mathcal{F}_n], n = 1, 2, \dots\}$ is a martingale. As $A_n, n = 1, 2, \dots$, is decreasing and $A_n \in \mathcal{F}_{n-1}$, then $\{Z_n, n = 1, 2, \dots\}$ is nonpositive submartingale since

$$\begin{aligned} E[Z_n | \mathcal{F}_{n-1}] &= E[E(A_\infty | \mathcal{F}_n) - A_n | \mathcal{F}_{n-1}] \\ &= E[A_\infty | \mathcal{F}_{n-1}] - A_n \\ &\geq E[A_\infty | \mathcal{F}_{n-1}] - A_{n-1} = Z_{n-1}. \end{aligned}$$

On the other hand, by monotone convergence theorem, (1.5.13) gives

$$\begin{aligned} \lim_{n \rightarrow \infty} E[Z_n] &= \lim_{n \rightarrow \infty} E[E(A_\infty | \mathcal{F}_n) - A_n] \\ &= E[A_\infty] - E[A_\infty] = 0. \end{aligned}$$

To show that such a composition is unique, suppose $X_n = Y'_n + Z'_n$ is another decomposition. Now let

$$W_n = Y_n - Y'_n = Z'_n - Z_n.$$

Then $\{W_n, n = 1, 2, \dots\}$ is a martingale and

$$\begin{aligned} E[|W_n|] &\leq E[|Z'_n|] + E[|Z_n|] \\ &= -E[Z'_n] - E[Z_n] \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

This implies that $\{W_n, n = 1, 2, \dots\}$ is uniformly integrable and $W_\infty = \lim_{n \rightarrow \infty} W_n = 0$. Consequently,

$$Y_n - Y'_n = W_n = \lim_{k \rightarrow \infty} E[W_{n+k} | \mathcal{F}_n] = E[W_\infty | \mathcal{F}_n] = 0.$$

Thus such a decomposition is unique. Now if in addition, $\{X_n, n = 1, 2, \dots\}$ are nonnegative, then (1.5.12) yields

$$\begin{aligned} Y_n &= M_n - E[A_\infty | \mathcal{F}_n] \\ &= \lim_{k \rightarrow \infty} E[M_{n+k} | \mathcal{F}_n] - \lim_{k \rightarrow \infty} E[A_{n+k} | \mathcal{F}_n] \\ &= \lim_{k \rightarrow \infty} E[M_{n+k} - A_{n+k} | \mathcal{F}_n] = \lim_{k \rightarrow \infty} E[X_{n+k} | \mathcal{F}_n] \geq 0. \end{aligned}$$

Thus, $\{Y_n, n = 1, 2, \dots\}$ are also nonnegative.

(2) \Rightarrow (1) Now

$$E[X_n] = E[Y_n] + E[Z_n] = E[Y_1] + E[Z_n] \leq E[Y_1] < \infty.$$

This completes the proof of Theorem 1.5.3.

Theorem 1.5.4. (The Krickeberg Decomposition)

Assume that $\{X_n, n = 1, 2, \dots\}$ is a submartingale with respect to $\{\mathcal{F}_n, n = 1, 2, \dots\}$. Then the following two conditions are equivalent.

$$(1) \quad \sup_n E[X_n^+] < \infty. \tag{1.5.14}$$

$$(2) \quad \{X_n, n = 1, 2, \dots\} \text{ has a decomposition such that} \\ X_n = L_n - M_n, \tag{1.5.15}$$

where $\{L_n, n = 1, 2, \dots\}$ is a nonpositive submartingale and $\{M_n, n = 1, 2, \dots\}$ is a nonpositive martingale. Moreover, such a decomposition has the maximality such that for any other decomposition $X_n = L'_n - M'_n$, where L'_n, M'_n are nonpositive submartingale and nonpositive martingale respectively, then

$$L_n \geq L'_n, \quad M_n \geq M'_n.$$

Proof.

(1) \Rightarrow (2) Because $\{X_n, n = 1, 2, \dots\}$ is a submartingale, then

$\{X_n^+, n = 1, 2, \dots\}$ is a nonnegative submartingale and $\{E[X_n^+], n = 1, 2, \dots\}$ is nondecreasing. As a result,

$$\lim_{n \rightarrow \infty} E[X_n^+] = \sup_n E[X_n^+] < \infty. \quad (1.5.16)$$

By Theorem 1.5.3, $\{X_n^+, n = 1, 2, \dots\}$ has a Riesz decomposition

$$X_n^+ = Y_n + Z_n,$$

where $\{Y_n, n = 1, 2, \dots\}$ is a nonnegative martingale and $\{Z_n, n = 1, 2, \dots\}$ is nonpositive submartingale. Now for $n = 1, 2, \dots$, let $M_n = -Y_n$ and define

$$L_n = X_n + M_n \leq X_n^+ + M_n = X_n^+ - Y_n = Z_n \leq 0.$$

Then $\{M_n, n = 1, 2, \dots\}$ is a nonpositive martingale and $\{L_n, n = 1, 2, \dots\}$ is a nonpositive submartingale. To show the maximality, suppose that $X_n = L'_n - M'_n$ is another decomposition, then

$$M'_n = L'_n - X_n \leq 0 \wedge (-X_n) = -X_n^+ = -Y_n - Z_n = M_n - Z_n.$$

Therefore

$$\begin{aligned} M'_n &= E[M'_{n+k} | \mathcal{F}_n] \\ &\leq E[-X_{n+k}^+ | \mathcal{F}_n] = E[M_{n+k} - Z_{n+k} | \mathcal{F}_n] \\ &= M_n - E[Z_{n+k} | \mathcal{F}_n]. \end{aligned} \quad (1.5.17)$$

Because $\{Z_n, n = 1, 2, \dots\}$ is nonpositive and

$$\lim_{n \rightarrow \infty} E[Z_n] = 0,$$

then $\{Z_n, n = 1, 2, \dots\}$ is uniformly integrable and $Z_\infty = \lim_{n \rightarrow \infty} Z_n = 0$. Thus, $\lim_{k \rightarrow \infty} E[Z_{n+k} | \mathcal{F}_n] = 0$. Consequently, letting $k \rightarrow \infty$, (1.5.17) yields

$$M'_n \leq M_n,$$

hence

$$L'_n - L_n = M'_n - M_n \leq 0.$$

(2) \Rightarrow (1) Since $M_n = L_n - X_n$, we have $M_n \leq -X_n^+$. Then

$$E[X_n^+] \leq -E[M_n] = -E[M_0] < \infty.$$

This completes the proof of Theorem 1.5.4.

1.6 * The Rate of Occurrence of Failures

Given a repairable system, let $N_f(t)$ be the number of failures of the system that have occurred by time t . Then the expected number of failures by time t is $M_f(t) = E[N_f(t)]$. Its derivative $m_f(t)$ is called the rate of occurrence of failures (ROCOF) at time t . Obviously, if $m_f(t)$ is increasing, the system is deteriorating, if $m_f(t)$ is decreasing, the system is improving. Therefore, the ROCOF is an important index in reliability theory.

Lam (1995, 1997) introduced a simple formula for the determination of the ROCOF $m_f(t)$ for a Markov chain with infinite state space. Suppose the state of a system at time t is $X(t)$. Assume that $\{X(t), t \geq 0\}$ is a continuous time homogeneous Markov chain with a finite or infinite state space $S = \{0, 1, 2, \dots\}$. Let the infinitesimal matrix of the process be $A = (q_{ij})$. Thus

$$\begin{aligned}
 p_{ij}(\Delta t) &= P(X(t + \Delta t) = j \mid X(t) = i) \\
 &= \begin{cases} q_{ij}\Delta t + o(\Delta t) & j \neq i, \\ 1 - q_i\Delta t + o(\Delta t) & j = i, \end{cases} \tag{1.6.1}
 \end{aligned}$$

where $q_i = -q_{ii} \geq 0$ and $q_{ij} \geq 0$ for $j \neq i$.

Assume that the system has two kinds of state: up state and down state say. Denote the set of up states by W and set of down states by F . Then obviously $S = W \cup F$. At the beginning, suppose the system is in an up state. Let the number of transitions of the Markov chain by time t be $N(t)$ and the number of failures by time t be $N_f(t)$ respectively. Denote $N(t, t + \Delta t] = N(t + \Delta t) - N(t)$ and $N_f(t, t + \Delta t] = N_f(t + \Delta t) - N_f(t)$. Then clearly

$$N_f(t, t + \Delta t] \leq N(t, t + \Delta t]. \tag{1.6.2}$$

Moreover, let $p_i(t) = P(X(t) = i)$. Now, we make the following two assumptions.

Assumption 1. The Markov chain $\{X(t), t \geq 0\}$ is conservative, i.e.

$$q_i = \sum_{j \neq i} q_{ij}.$$

Assumption 2.

$$q = \sup_i q_i < \infty.$$

Note that if the Markov chain has a finite state space S , Assumptions 1 and 2 will automatically hold. Thus Assumptions 1 and 2 are actually made for

the Markov chain with an infinite state space S .

It is well known that for a Markov chain $\{X(t), t \geq 0\}$, the sojourn time at state i has an exponential distribution $\text{Exp}(q_i)$. Therefore, Assumption 2 means that the expected sojourn time will have a positive lower bound, i.e.

$$\frac{1}{q} = \inf_i \frac{1}{q_i} > 0.$$

To derive the formula for the determination of the ROCOF $m_f(t)$ for the Markov chain $\{X(t), t \geq 0\}$, we start with time $T_0 = 0$, for $n > 1$, let T_n be the n th transition time of the process $\{X(t), t \geq 0\}$. If $X(T_n) = i$, then $T_{n+1} - T_n$ is the sojourn time of the process in state i . Now we can prove the following lemmas.

Lemma 1.6.1. For all integer n , we have

$$1. P(T_n - T_0 \leq \Delta t \mid X(T_n) = j, X(T_0) = i) \leq (q\Delta t)^n, \quad (1.6.3)$$

$$2. P(T_n - T_0 \leq \Delta t \mid X(T_0) = i) \leq (q\Delta t)^n. \quad (1.6.4)$$

Proof.

For $n = 1$, it is well known that the sojourn time in state i has exponential distribution $\text{Exp}(q_i)$ with density function $f(x) = q_i \exp(-q_i x)$ for $x > 0$ and 0 otherwise, and T_1 is independent of the state $X(T_1) = j$. Therefore

$$\begin{aligned} &P(T_1 - T_0 \leq \Delta t \mid X(T_1) = j, X(T_0) = i) \\ &= P(T_1 - T_0 \leq \Delta t \mid X(T_0) = i) = 1 - \exp(-q_i \Delta t) \\ &\leq q_i \Delta t \leq q \Delta t, \end{aligned} \quad (1.6.5)$$

where (1.6.5) is due to the inequality $1 - \exp(-x) \leq x$ for $x \geq 0$ and Assumption 2. Thus, (1.6.3) and (1.6.4) are true for $n = 1$. Now, assume that (1.6.3) and (1.6.4) hold for n . For $n + 1$,

$$\begin{aligned} &P(T_{n+1} - T_0 \leq \Delta t \mid X(T_{n+1}) = j, X(T_0) = i) \\ &\leq \sum_k P(T_{n+1} - T_n \leq \Delta t, T_n - T_0 \leq \Delta t, X(T_n) = k \\ &\quad \mid X(T_{n+1}) = j, X(T_0) = i) \\ &= \sum_k P(T_{n+1} - T_n \leq \Delta t \\ &\quad \mid T_n - T_0 \leq \Delta t, X(T_n) = k, X(T_{n+1}) = j, X(T_0) = i) \\ &\quad \times P(T_n - T_0 \leq \Delta t \mid X(T_n) = k, X(T_{n+1}) = j, X(T_0) = i) \\ &\quad \times P(X(T_n) = k \mid X(T_{n+1}) = j, X(T_0) = i). \end{aligned} \quad (1.6.6)$$

Then

$$\begin{aligned}
 & P(T_{n+1} - T_0 \leq \Delta t \mid X(T_{n+1}) = j, X(T_0) = i) \\
 & \leq \sum_k P(T_{n+1} - T_n \leq \Delta t \mid X(T_{n+1}) = j, X(T_n) = k) \\
 & \quad \times P(T_n - T_0 \leq \Delta t \mid X(T_n) = k, X(T_0) = i) \\
 & \quad \times P(X(T_n) = k \mid X(T_{n+1}) = j, X(T_0) = i) \tag{1.6.7} \\
 & \leq (q\Delta t)^{n+1} \sum_k P(X(T_n) = k \mid X(T_{n+1}) = j, X(T_0) = i) \tag{1.6.8} \\
 & = (q\Delta t)^{n+1},
 \end{aligned}$$

where (1.6.7) is because of the Markov property and (1.6.8) is due to the homogeneity. Consequently by induction, (1.6.3) holds for all integers. A similar argument shows that (1.6.4) is also true for all integers.

Lemma 1.6.2.

$$P(T_1 - T_0 \leq \Delta t, T_2 - T_0 \geq \Delta t, X(T_1) = j \mid X(T_0) = i) = q_{ij}\Delta t + o(\Delta t). \tag{1.6.9}$$

Proof.

$$\begin{aligned}
 & P(T_1 - T_0 \leq \Delta t, T_2 - T_0 \geq \Delta t, X(T_1) = j \mid X(T_0) = i) \\
 & = \int_0^\infty P(T_1 - T_0 \leq \Delta t, T_2 - T_0 > \Delta t, X(T_1) = j \mid T_1 - T_0 = x, X(T_0) = i) \\
 & \quad \times q_i \exp(-q_i x) dx \\
 & = \int_0^{\Delta t} P(T_2 - T_1 > \Delta t - x \mid X(T_1) = j, T_1 - T_0 = x, X(T_0) = i) \\
 & \quad \times P(X(T_1) = j \mid T_1 - T_0 = x, X(T_0) = i) q_i \exp(-q_i x) dx \\
 & = \int_0^{\Delta t} P(T_2 - T_1 > \Delta t - x \mid X(T_1) = j) [(q_{ij}/q_i) + o(1)] q_i \exp(-q_i x) dx \\
 & = \int_0^{\Delta t} \exp[-q_j(\Delta t - x)] [(q_{ij}/q_i) + o(1)] q_i \exp(-q_i x) dx \\
 & = q_{ij}\Delta t + o(\Delta t).
 \end{aligned}$$

Lemma 1.6.3.

$$P(N_f(t, t + \Delta t] = 1) = \sum_{i \in W, j \in F} p_i(t) q_{ij} \Delta t + o(\Delta t).$$

Proof.

Because of (1.6.2), we have

$$P(N_f(t, t + \Delta t] = 1)$$

$$\begin{aligned}
&= \sum_{i \in W} P(N_f(t, t + \Delta t] = 1 \mid X(t) = i)P(X(t) = i) \\
&\quad + \sum_{i \in F} P(N_f(t, t + \Delta t] = 1 \mid X(t) = i)P(X(t) = i) \\
&= \sum_{i \in W} \{P(N_f(t, t + \Delta t] = 1, N(t, t + \Delta t] = 1 \mid X(t) = i)P(X(t) = i) \\
&\quad + P(N_f(t, t + \Delta t] = 1, N(t, t + \Delta t] \geq 2 \mid X(t) = i)P(X(t) = i)\} \\
&\quad + \sum_{i \in F} P(N_f(t, t + \Delta t] = 1, N(t, t + \Delta t] \geq 2 \mid X(t) = i)P(X(t) = i) \\
&= \sum_{i \in W} P(N_f(t, t + \Delta t] = 1, N(t, t + \Delta t] = 1 \mid X(t) = i)P(X(t) = i) \\
&\quad + \sum_i P(N_f(t, t + \Delta t] = 1, N(t, t + \Delta t] \geq 2 \mid X(t) = i)P(X(t) = i) \\
&= I_1 + I_2. \tag{1.6.10}
\end{aligned}$$

On the one hand, it follows from Lemma 1.6.2 that

$$\begin{aligned}
I_1 &= \sum_{i \in W, j \in F} P(T_1 - T_0 \leq \Delta t, T_2 - T_0 \geq \Delta t, X(T_1) = j \mid X(T_0) = i)p_i(t) \\
&= \sum_{i \in W, j \in F} p_i(t)q_{ij}\Delta t + o(\Delta t). \tag{1.6.11}
\end{aligned}$$

On the other hand, Lemma 1.6.1 gives

$$\begin{aligned}
I_2 &\leq \sum_i P(N(t, t + \Delta t] \geq 2 \mid X(T_0) = i)p_i(t) \\
&= \sum_i P(T_2 - T_0 \leq \Delta t \mid X(T_0) = i)p_i(t) \\
&\leq (q\Delta t)^2 \sum_i p_i(t) = (q\Delta t)^2. \tag{1.6.12}
\end{aligned}$$

The combination of (1.6.11) and (1.6.12) yields Lemma 1.6.3.

Lemma 1.6.4.

$$\sum_{k=2}^{\infty} kP(N_f(t, t + \Delta t] = k) = o(\Delta t).$$

Proof.

$$\begin{aligned}
 & \sum_{k=2}^{\infty} kP(N_f(t, t + \Delta t] = k) \\
 = & \sum_{k=2}^{\infty} \sum_i kP(N_f(t, t + \Delta t] = k \mid X(t) = i)P(X(t) = i) \\
 \leq & \sum_{k=2}^{\infty} \sum_i kP(N(t, t + \Delta t] \geq k \mid X(t) = i)P(X(t) = i) \\
 = & \sum_{k=2}^{\infty} \sum_i kP(T_k - T_0 \leq \Delta t \mid X(T_0) = i)P(X(t) = i) \\
 \leq & \sum_{k=2}^{\infty} k(q\Delta t)^k = o(\Delta t),
 \end{aligned} \tag{1.6.13}$$

where (1.6.13) follows from Lemma 1.6.1.

Now, since

$$\begin{aligned}
 M_f(t + \Delta t) - M_f(t) &= E[N_f(t, t + \Delta T)] \\
 &= P(N_f(t, t + \Delta t] = 1) + \sum_{k=2}^{\infty} kP(N_f(t, t + \Delta t] = k) \\
 &= \sum_{i \in W, j \in F} p_i(t)q_{ij}\Delta t + o(\Delta t),
 \end{aligned}$$

by Lemmas 1.6.3 and 1.6.4. Thus we have proven the following theorem.

Theorem 1.6.5. Assume that a Markov chain is conservative, i.e.

$$q_i = \sum_{j \neq i} q_{ij}$$

and

$$q = \sup_i q_i < \infty,$$

then the ROCOF at time t is given by

$$m_f(t) = \sum_{i \in W, j \in F} p_i(t)q_{ij}. \tag{1.6.14}$$

In particular, if the state space S of a Markov chain is finite, Assumptions 1 and 2 are clearly true, then Theorem 1.6.5 holds. This special case was considered by Shi (1985).

1.7 Notes and References

For reading this monograph, we assume that the readers should have learned Calculus and an undergraduate course in Probability and Statistics. Of course, some other preliminaries are needed. Chapter 1 contains almost all additional knowledge needed in this monograph. Sections 1.2 to 1.4 are mainly based on Ross (1996), Barlow and Proschan (1975), Ascher and Feingold (1984). However, in Section 1.4, the class of life distributions ERBLE (ERWLE) that takes into account the effect of preventive repair is new, it was introduced by Lam (2007a). We expect that ERBLE (ERWLE) will have important theoretical interest and wide practical application. In Section 1.5, we introduce martingales and the martingale convergence theorem. Moreover, the Doob, Riesz and Krickberg decomposition theorems are also studied. For reference of these three decomposition theorems, see Dellacherie and Meyer (1982) or He et al. (1995). As in many reference books, they just highlight on the supermartingale case. Here we state the martingale convergence theorem for both the supermartingale and submartingale cases. Then we state these three decomposition theorems for submartingale case and give detailed proof for convenience of application in geometric process. In Section 1.6, we study the rate of occurrence of failures (ROCOF). If the expected sojourn time in a state of a conservative continuous time Markov chain has a positive lower bound, Lam (1995, 1997) gave a formula for the evaluation of ROCOF. In the case that a process is not a Markov chain but can be reduced to a Markov process after introducing some supplementary variables, Lam (1995, 1997) also gave a formula for the determination of the ROCOF. Section 1.6 is based on Lam (1997) that was published in *Journal of Applied Probability* by Applied Probability Trust.