2

System Reliability Concepts

The analysis of the reliability of a system must be based on precisely defined concepts. Since it is readily accepted that a population of supposedly identical systems, operating under similar conditions, fall at different points in time, then a failure phenomenon can only be described in probabilistic terms. Thus, the fundamental definitions of reliability must depend on concepts from probability theory. This chapter describes the concepts of system reliability engineering. These concepts provide the basis for quantifying the reliability of a system. They allow precise comparisons between systems or provide a logical basis for improvement in a failure rate. Various examples reinforce the definitions as presented in Section 2.1. Section 2.2 examines common distribution functions useful in reliability engineering. Several distribution models are discussed and the resulting hazard functions are derived. Section 2.3 describes a new concept of systemability. Several systemability functions of various system configurations such as series, parallel, and k-out-of-n, are presented. Section 2.4 discusses various reliability aspects of systems with multiple failure modes. Stochastic processes including Markov process, Poisson process, renewal process, quasi-renewal process, and nonhomogeneous Poisson process are discussed in Sections 2.5 and 2.6.

In general, a system may be required to perform various functions, each of which may have a different reliability. In addition, at different times, the system may have a different probability of successfully performing the required function under stated conditions. The term failure means that the system is not capable of performing a function when required. The term *capable* used here is to define if the system is capable of performing the required function. However, the term *capable* is unclear and only various degrees of capability can be defined.

2.1 Reliability Measures

The reliability definitions given in the literature vary between different practitioners as well as researchers. The generally accepted definition is as follows.

Definition 2.1: *Reliability* is the probability of success or the probability that the system will perform its intended function under specified design limits.

More specific, reliability is the probability that a product or part will operate properly for a specified period of time (design life) under the design operating conditions (such as temperature, volt, *etc.*) without failure. In other words, reliability may be used as a measure of the system's success in providing its function properly. Reliability is one of the quality characteristics that consumers require from the manufacturer of products.

Mathematically, reliability $R(t)$ *is* the probability that a system will be successful in the interval from time 0 to time t:

$$R(t) = P(T > t) \qquad t \geq 0 \tag{2.1}$$

where T is a random variable denoting the time-to-failure or failure time.

Unreliability $F(t)$, a measure of failure, is defined as the probability that the system will fail by time t:

$$F(t) = P(T \leq t) \qquad \text{for } t \geq 0$$

In other words, $F(t)$ *is* the failure distribution function. If the time-to-failure random variable T has a density function $f(t)$, then

$$R(t) = \int_t^\infty f(s)ds$$

or, equivalently,

$$f(t) = -\frac{d}{dt}[R(t)]$$

The density function can be mathematically described in terms of T:

$$\lim_{\Delta t \to 0} P(t < T \leq t + \Delta t)$$

This can be interpreted as the probability that the failure time T will occur between the operating time t and the next interval of operation, $t+\Delta t$.

Consider a new and successfully tested system that operates well when put into service at time $t = 0$. The system becomes less likely to remain successful as the time interval increases. The probability of success for an infinite time interval, of course, is zero.

Thus, the system functions at a probability of one and eventually decreases to a probability of zero. Clearly, reliability is a function of mission time. For example, one can say that the reliability of the system is 0.995 for a mission time of 24 hours. However, a statement such as the reliability of the system is 0.995 is meaningless because the time interval is unknown.

Example 2.1: A computer system has an exponential failure time density function

$$f(t) = \frac{1}{9{,}000} e^{-\frac{t}{9{,}000}} \qquad t \geq 0$$

What is the probability that the system will fail after the warranty (six months or 4380 hours) and before the end of the first year (one year or 8760 hours)?

Solution: From equation (2.1) we obtain

$$P(4380 < T \leq 8760) = \int_{4380}^{8760} \frac{1}{9000} e^{-\frac{t}{9000}} dt$$
$$= 0.237$$

This indicates that the probability of failure during the interval from six months to one year is 23.7%.

If the time to failure is described by an exponential failure time density function, then

$$f(t) = \frac{1}{\theta} e^{-\frac{t}{\theta}} \qquad t \geq 0,\ \theta > 0 \tag{2.2}$$

and this will lead to the reliability function

$$R(t) = \int_{t}^{\infty} \frac{1}{\theta} e^{-\frac{s}{\theta}} ds = e^{-\frac{t}{\theta}} \qquad t \geq 0 \tag{2.3}$$

Consider the Weibull distribution where the failure time density function is given by

$$f(t) = \frac{\beta t^{\beta-1}}{\theta^{\beta}} e^{-\left(\frac{t}{\theta}\right)^{\beta}} \qquad t \geq 0,\ \theta > 0,\ \beta > 0$$

Then the reliability function is

$$R\ (t)\ =\ e^{-\left(\frac{t}{\theta}\right)^{\beta}} \qquad t \geq 0$$

Thus, given a particular failure time density function or failure time distribution function, the reliability function can be obtained directly. Section 2.2 provides further insight for specific distributions.

System Mean Time to Failure

Suppose that the reliability function for a system is given by $R(t)$. The expected failure time during which a component is expected to perform successfully, or the system mean time to failure (MTTF), is given by

$$MTTF = \int_{0}^{\infty} t\, f(t) dt \tag{2.4}$$

Substituting

$$f(t) = -\frac{d}{dt}[R(t)]$$

into equation (2.4) and performing integration by parts, we obtain

$$
\begin{aligned}
MTTF &= -\int_0^\infty t d[R(t)] \\
&= [-tR(t)]\Big|_0^\infty + \int_0^\infty R(t)dt
\end{aligned}
\tag{2.5}
$$

The first term on the right-hand side of equation (2.5) equals zero at both limits, since the system must fail after a finite amount of operating time. Therefore, we must have $tR(t) \to 0$ as $t \to \infty$. This leaves the second term, which equals

$$
MTTF = \int_0^\infty R(t)dt \tag{2.6}
$$

Thus, MTTF is the definite integral evaluation of the reliability function. In general, if $\lambda(t)$ is defined as the failure rate function, then, by definition, MTTF is not equal to $1/\lambda(t)$.

The MTTF should be used when the failure time distribution function is specified because the reliability level implied by the MTTF depends on the underlying failure time distribution. Although the MTTF measure is one of the most widely used reliability calculations, it is also one of the most misused calculations. It has been misinterpreted as "guaranteed minimum lifetime". Consider the results as given in Table 2.1 for a twelve-component life duration test.

Table 2.1. Results of a twelve-component life duration test

Component	Time to failure (hours)
1	4510
2	3690
3	3550
4	5280
5	2595
6	3690
7	920
8	3890
9	4320
10	4770
11	3955
12	2750

Using a basic averaging technique, the component MTTF of 3660 hours was estimated. However, one of the components failed after 920 hours. Therefore, it is important to note that the system MTTF denotes the average time to failure. It is neither the failure time that could be expected 50% of the time, nor is it the guaranteed minimum time of system failure.

A careful examination of equation (2.6) will show that two failure distributions can have the same MTTF and yet produce different reliability levels. This is illustrated in a case where the MTTFs are equal, but with normal and exponential failure distributions. The normal failure distribution is symmetrical about its mean, thus

$$R(MTTF) = P(Z \geq 0) = 0.5$$

where Z is a standard normal random variable. When we compute for the exponential failure distribution using equation (2.3), recognizing that θ = MTTF, the reliability at the MTTF is

$$R(MTTF) = e^{-\frac{MTTF}{MTTF}} = 0.368$$

Clearly, the reliability in the case of the exponential distribution is about 74% of that for the normal failure distribution with the same MTTF.

Failure Rate Function

The probability of a system failure in a given time interval [t_1, t_2] can be expressed in terms of the reliability function as

$$\int_{t_1}^{t_2} f(t)dt = \int_{t_1}^{\infty} f(t)dt - \int_{t_2}^{\infty} f(t)dt$$
$$= R(t_1) - R(t_2)$$

or in terms of the failure distribution function (or the unreliability function) as

$$\int_{t_1}^{t_2} f(t)dt = \int_{-\infty}^{t_2} f(t)dt - \int_{-\infty}^{t_1} f(t)dt$$
$$= F(t_2) - F(t_1)$$

The rate at which failures occur in a certain time interval [t_1, t_2] *is* called the *failure rate.* It is defined as the probability that a failure per unit time occurs in the interval, given that a failure has not occurred prior to t_1, the beginning of the interval. Thus, the failure rate is

$$\frac{R(t_1) - R(t_2)}{(t_2 - t_1)R(t_1)}$$

Note that the failure rate is a function of time. If we redefine the interval as $[t, t + \Delta t]$, the above expression becomes

$$\frac{R(t) - R(t + \Delta t)}{\Delta t R(t)}$$

The rate in the above definitions is expressed as failures per unit time, when in reality, the time units might be in terms of miles, hours, etc. The *hazard function is* defined as the limit of the failure rate as the interval approaches zero. Thus, the hazard function *h(t) is* the instantaneous failure rate, and is defined by

$$h(t) = \lim_{\Delta t \to 0} \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)}$$
$$= \frac{1}{R(t)}[-\frac{d}{dt}R(t)] \qquad (2.7)$$
$$= \frac{f(t)}{R(t)}$$

The quantity $h(t)dt$ represents the probability that a device of age t will fail in the small interval of time t to $(t+dt)$. The importance of the hazard function is that it indicates the change in the failure rate over the life of a population of components by plotting their hazard functions on a single axis. For example, two designs may provide the same reliability at a specific point in time, but the failure rates up to this point in time can differ.

The death rate, in statistical theory, is analogous to the failure rate as the force of mortality is analogous to the hazard function. Therefore, the hazard function or hazard rate or failure rate function is the ratio of the probability density function (pdf) to the reliability function.

Maintainability

When a system fails to perform satisfactorily, repair is normally carried out to locate and correct the fault. The system is restored to operational effectiveness by making an adjustment or by replacing a component.

Maintainability is defined as the probability that a failed system will be restored to specified conditions within a given period of time when maintenance is performed according to prescribed procedures and resources. In other words, maintainability is the probability of isolating and repairing a fault in a system within a given time. Maintainability engineers must work with system designers to ensure that the system product can be maintained by the customer efficiently and cost effectively. This function requires the analysis of part removal, replacement, tear-down, and build-up of the product in order to determine the required time to carry out the operation, the necessary skill, the type of support equipment and the documentation.

Let T denote the random variable of the time to repair or the total downtime. If the repair time T has a repair time density function $g(t)$, then the maintainability, $V(t)$, is defined as the probability that the failed system will be back in service by time t, *i.e.*,

$$V(t) = P(T \le t) = \int_0^t g(s)ds$$

For example, if $g(t) = \mu e^{-\mu t}$ where $\mu > 0$ is a constant repair rate, then

$$V(t) = 1 - e^{-\mu t}$$

which represents the exponential form of the maintainability function.

An important measure often used in maintenance studies is the mean time to repair (MTTR) or the mean downtime. MTTR is the expected value of the random variable repair time, not failure time, and is given by

$$MTTR = \int_0^\infty t g(t)dt$$

When the distribution has a repair time density given by $g(t) = \mu e^{-\mu t}$, then, from the above equation, MTTR = $1/\mu$. When the repair time T has the log normal density function $g(t)$, and the density function is given by

$$g(t) = \frac{1}{\sqrt{2\pi}\,\sigma t} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} \qquad t > 0$$

then it can be shown that

$$MTTR = m\, e^{\frac{\sigma^2}{2}}$$

where m denotes the median of the log normal distribution.

In order to design and manufacture a maintainable system, it is necessary to predict the MTTR for various fault conditions that could occur in the system. This is generally based on past experiences of designers and the expertise available to handle repair work.

The system repair time consists of two separate intervals: passive repair time and active repair time. Passive repair time is mainly determined by the time taken by service engineers to travel to the customer site. In many cases, the cost of travel time exceeds the cost of the actual repair. Active repair time is directly affected by the system design and is listed as follows:

1. The time between the occurrence of a failure and the system user becoming aware that it has occurred.
2. The time needed to detect a fault and isolate the replaceable component(s).
3. The time needed to replace the faulty component(s).
4. The time needed to verify that the fault has been corrected and the system is fully operational.

The active repair time can be improved significantly by designing the system in such a way that faults may be quickly detected and isolated. As more complex systems are designed, it becomes more difficult to isolate the faults.

Availability

Reliability is a measure that requires system success for an entire mission time. No failures or repairs are allowed. Space missions and aircraft flights are examples of systems where failures or repairs are not allowed. Availability is a measure that allows for a system to repair when failure occurs.

The availability of a system is defined as the probability that the system is successful at time t. Mathematically,

$$\text{Availability} = \frac{\text{System up time}}{\text{System up time} + \text{System down time}}$$

$$= \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}$$

Availability is a measure of success used primarily for repairable systems. For non-repairable systems, availability, $A(t)$, equals reliability, $R(t)$. In repairable systems, $A(t)$ will be equal to or greater than $R(t)$.

The mean time between failures (MTBF) is an important measure in repairable systems. This implies that the system has failed and has been repaired. Like MTTF and MTTR, MTBF is an expected value of the random variable time between failures. Mathematically,

$$\text{MTBF} = \text{MTTF} + \text{MTTR}$$

The term MTBF has been widely misused. In practice, MTTR is much smaller than MTTF, which is approximately equal to MTBF. MTBF is often incorrectly substituted for MTTF, which applies to both repairable systems and non-repairable systems. If the MTTR can be reduced, availability will increase, and the system will be more economical.

A system where faults are rapidly diagnosed is more desirable than a system that has a lower failure rate but where the cause of a failure takes longer to detect, resulting in a lengthy system downtime. When the system being tested is renewed through maintenance and repairs, $E(T)$ is also known as MTBF.

2.2 Common Distribution Functions

This section presents some of the common distribution functions and several hazard models that have applications in reliability engineering (Pham 2000a).

Binomial Distribution

The binomial distribution is one of the most widely used discrete random variable distributions in reliability and quality inspection. It has applications in reliability engineering, *e.g.,* when one is dealing with a situation in which an event is either a success or a failure.

The pdf of the distribution is given by

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, ..., n$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

where n = number of trials; x = number of successes; p = single trial probability of success.

The reliability function, $R(k)$, (*i.e.*, at least k out of n items are good) is given by

$$R(k) = \sum_{x=k}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$

Example 2.2: Suppose in the production of lightbulbs, 90% are good. In a random sample of 20 lightbulbs, what is the probability of obtaining at least 18 good lightbulbs?

Solution: The probability of obtaining 18 or more good lightbulbs in the sample of 20 is

$$R(18)=\sum_{x=18}^{20}\binom{20}{18}(0.9)^x(0.1)^{20-x}$$
$$=0.667$$

Poisson Distribution

Although the Poisson distribution can be used in a manner similar to the binomial distribution, it is used to deal with events in which the sample size is unknown. This is also a discrete random variable distribution whose pdf is given by

$$P(X=x)=\frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad \text{for } x = 0, 1, 2, ...$$

where λ = constant failure rate, x = is the number of events. In other words, $P(X = x)$ is the probability of exactly x failures occurring in time t. Therefore, the reliability Poisson distribution, $R(k)$ (the probability of k or fewer failures) is given by

$$R(k)=\sum_{x=0}^{k}\frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

This distribution can be used to determine the number of spares required for the reliability of standby redundant systems during a given mission.

Exponential Distribution

Exponential distribution plays an essential role in reliability engineering because it has a constant failure rate. This distribution has been used to model the lifetime of electronic and electrical components and systems. This distribution is appropriate when a used component that has not failed is as good as a new component - a rather restrictive assumption. Therefore, it must be used diplomatically since numerous applications exist where the restriction of the memoryless property may not apply. For this distribution, we have reproduced equations (2.2) and (2.3), respectively:

$$f(t)=\frac{1}{\theta}e^{-\frac{t}{\theta}}=\lambda e^{-\lambda t}, \quad t\ge 0$$
$$R(t)=e^{-\frac{t}{\theta}}=e^{-\lambda t}, \quad t\ge 0$$

where $\theta = 1/\lambda > 0$ is an MTTF's parameter and $\lambda \ge 0$ is a constant failure rate.

The hazard function or failure rate for the exponential density function is constant, *i.e.*,

$$h(t)=\frac{f(t)}{R(t)}=\frac{\frac{1}{\theta}e^{-\frac{1}{\theta}}}{e^{-\frac{1}{\theta}}}=\frac{1}{\theta}=\lambda$$

The failure rate for this distribution is λ, a constant, which is the main reason for this widely used distribution. Because of its constant failure rate property, the exponential is an excellent model for the long flat "intrinsic failure" portion of the bathtub curve. Since most parts and systems spend most of their lifetimes in this portion of the bathtub curve, this justifies frequent use of the exponential (when early failures or wear out is not a concern). The exponential model works well for inter-arrival times. When these events trigger failures, the exponential lifetime model can be used.

We will now discuss some properties of the exponential distribution that are useful in understanding its characteristics, when and where it can be applied.

Property 2.1: (*Memoryless property*) The exponential distribution is the only continuous distribution satisfying

$$P\{T \geq t\} = P\{T \geq t+s \mid T \geq s\} \quad \text{for } t > 0,\ s > 0 \tag{2.8}$$

This result indicates that the conditional reliability function for the lifetime of a component that has survived to time s is identical to that of a new component. This term is the so-called "used-as-good-as-new" assumption.

The lifetime of a fuse in an electrical distribution system may be assumed to have an exponential distribution. It will fail when there is a power surge causing the fuse to burn out. Assuming that the fuse does not undergo any degradation over time and that power surges that cause failure are likely to occur equally over time, then use of the exponential lifetime distribution is appropriate, and a used fuse that has not failed is as good as new.

Property 2.2: If T_1, T_2, ..., T_n, are independently and identically distributed exponential random variables (RVs) with a constant failure rate λ, then

$$2\lambda \sum_{i=1}^{n} T_i \sim \chi^2(2n)$$

where $\chi^2(2n)$ is a chi-squared distribution with degrees of freedom $2n$. This result is useful for establishing a confidence interval for λ.

Example 2.3: A manufacturer performs an operational life test on ceramic capacitors and finds they exhibit constant failure rate with a value of $3\text{x}10^{-8}$ failure per hour. What is the reliability of a capacitor at 10^4 hours?

Solution: The reliability of a capacitor at 10^4 hour is

$$R(t) = e^{-\lambda t} = e^{-3\times10^{-8}t} = e^{-3\times10^{-4}} = 0.9997$$

The resulting reliability plot is shown in Figure 2.1.

Normal Distribution

Normal distribution plays an important role in classical statistics owing to the *Central Limit Theorem*. In reliability engineering, the normal distribution primarily applies to measurements of product susceptibility and external stress. This two-parameter distribution is used to describe systems in which a failure results due to some wearout effect for many mechanical systems.

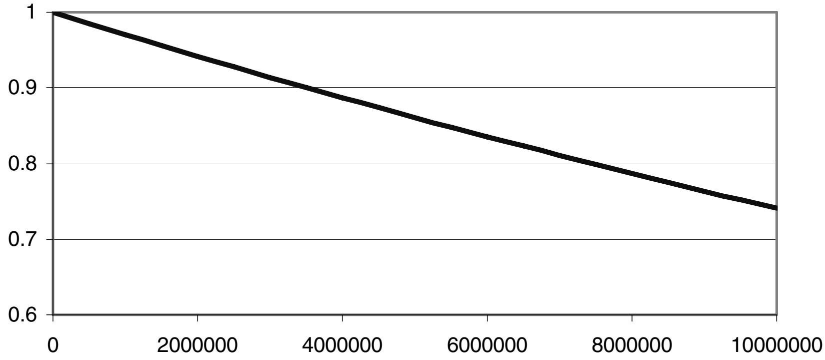

Figure 2.1. Reliability function *vs* time

The normal distribution takes the well-known bell shape. This distribution is symmetrical about the mean and the spread is measured by variance. The larger the value, the flatter the distribution. The pdf is given by

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} \qquad -\infty < t < \infty$$

where μ is the mean value and σ is the standard deviation. The cumulative distribution function (cdf) is

$$F(t) = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{s-\mu}{\sigma})^2} ds$$

The reliability function is

$$R(t) = \int_{t}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{s-\mu}{\sigma})^2} ds$$

There is no closed form solution for the above equation. However, tables for the standard normal density function are readily available (see Table A1.1 in Appendix 1) and can be used to find probabilities for any normal distribution. If

$$Z = \frac{T-\mu}{\sigma}$$

is substituted into the normal pdf, we obtain

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad -\infty < Z < \infty$$

This is a so-called standard normal pdf, with a mean value of *0* and a standard deviation of 1. The standardized cdf is given by

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds \tag{2.9}$$

where Φ is a standard normal distribution function. Thus, for a normal random variable T, with mean μ and standard deviation σ,

$$P(T \leq t) = P\left(Z \leq \frac{t-\mu}{\sigma}\right) = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

where Φ yields the relationship necessary if standard normal tables are to be used. The hazard function for a normal distribution is a monotonically increasing function of t. This can be easily shown by proving that $h'(t) \geq 0$ for all t. Since

$$h(t) = \frac{f(t)}{R(t)}$$

then (see Problem 15)

$$h'(t) = \frac{R(t)f'(t) + f^2(t)}{R^2(t)} \geq 0 \tag{2.10}$$

One can try this proof by employing the basic definition of a normal density function f.

Example 2.4: A component has a normal distribution of failure times with μ = 2000 hours and σ = 100 hours. Find the reliability of the component and the hazard function at 1900 hours.

Solution: The reliability function is related to the standard normal deviate z by

$$R(t) = P\left(Z > \frac{t-\mu}{\sigma}\right)$$

where the distribution function for Z is given by equation (2.9). For this particular application,

$$R(1900) = P\left(Z > \frac{1900-2000}{100}\right)$$
$$= P\left(z > -1\right)$$

From the standard normal table in Table A1.1 in Appendix 1, we obtain

$$R(1900) = 1 - \Phi(-1) = 0.8413.$$

The value of the hazard function is found from the relationship

$$h(t) = \frac{f(t)}{R(t)} = \frac{\Phi\left(z = \frac{t-\mu}{\sigma}\right)}{\sigma R(t)}$$

where ϕ is a pdf of standard normal density. Here

$$h(1900) = \frac{\Phi(-1.0)}{\sigma R(t)} = \frac{0.1587}{100(0.8413)}$$
$$= 0.0019 \text{ failures/cycle}$$

Example 2.5: A part has a normal distribution of failure times with μ = 40000 cycles and σ = 2000 cycles. Find the reliability of the part at 38000 cycles.

Solution: The reliability at 38000 cycles

$$R(38000) = P\left(z > \frac{38000-40000}{2000}\right)$$
$$= P(z > -1.0)$$
$$= \Phi(1.0) = 0.8413$$

The resulting reliability plot is shown in Figure 2.2.

The normal distribution is flexible enough to make it a very useful empirical model. It can be theoretically derived under assumptions matching many failure mechanisms. Some of these are corrosion, migration, crack growth, and in general, failures resulting from chemical reactions or processes. That does not mean that the normal is always the correct model for these mechanisms, but it does perhaps explain why it has been empirically successful in so many of these cases.

Log Normal Distribution

The log normal lifetime distribution is a very flexible model that can empirically fit many types of failure data. This distribution, with its applications in maintainability engineering, is able to model failure probabilities of repairable systems and to model the uncertainty in failure rate information. The log normal density function is given by

$$f(t) = \frac{1}{\sigma t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2} \qquad t \geq 0 \tag{2.11}$$

where μ and σ are parameters such that $-\infty < \mu < \infty$, and $\sigma > 0$. Note that μ and σ are not the mean and standard deviations of the distribution.

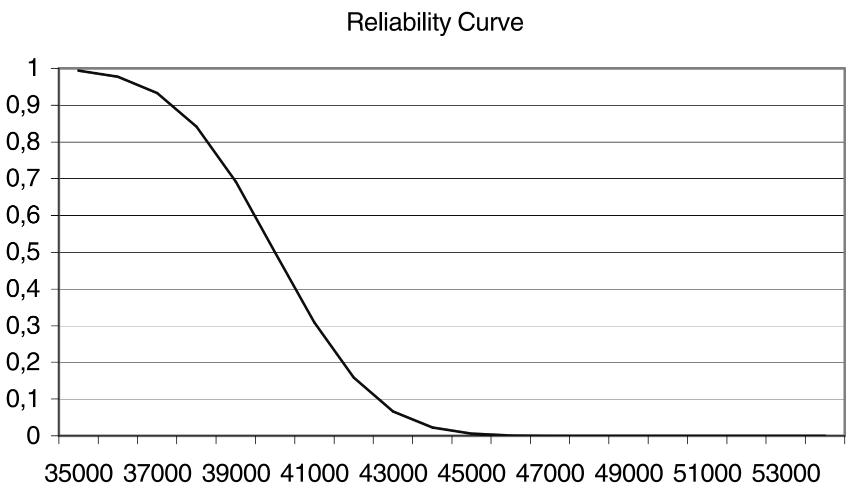

Figure 2.2. Normal reliability plot *vs* time

The relationship to the normal (just take natural logarithms of all the data and time points and you have "normal" data) makes it easy to work with many good software analysis programs available to treat normal data.

Mathematically, if a random variable X is defined as $X = lnT$, then X is normally distributed with a mean of μ and a variance of σ^2. That is,

$$E(X) = E(lnT) = \mu$$

and

$$V(X) = V(lnT) = \sigma^2.$$

Since $T = e^X$, the mean of the log normal distribution can be found by using the normal distribution. Consider that

$$E(T) = E(e^X) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\left[x - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]} dx$$

and by rearrangement of the exponent, this integral becomes

$$E(T) = e^{\mu+\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}[x-(\mu+\sigma^2)]^2} dx$$

Thus, the mean of the log normal distribution is

$$E(T) = e^{\mu+\frac{\sigma^2}{2}}$$

Proceeding in a similar manner,

$$E(T^2) = E(e^{2X}) = e^{2(\mu+\sigma^2)}$$

thus, the variance for the log normal is

$$V(T) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

The cumulative distribution function for the log normal is

$$F(t) = \int_0^t \frac{1}{\sigma s\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\ln s-\mu}{\sigma})^2} ds$$

and this can be related to the standard normal deviate Z by

$$F(t) = P[T \le t] = P(\ln T \le \ln t)$$
$$= P\left[Z \le \frac{\ln t - \mu}{\sigma}\right]$$

Therefore, the reliability function is given by

$$R(t) = P\left[Z > \frac{\ln t - \mu}{\sigma}\right] \tag{2.12}$$

and the hazard function would be

$$h(t) = \frac{f(t)}{R(t)} = \frac{\Phi\left(\frac{\ln t-\mu}{\sigma}\right)}{\sigma t\, R(t)}$$

where Φ is a cdf of standard normal density.

Example 2.6: The failure time of a certain component is log normal distributed with $\mu = 5$ and $\sigma = 1$. Find the reliability of the component and the hazard rate for a life of 50 time units.

Solution: Substituting the numerical values of μ, σ, and t into equation (2.12), we compute

$$R(50) = P\left[Z > \frac{\ln 50 - 5}{1}\right] = P[Z > -1.09]$$
$$= 0.8621$$

Similarly, the hazard function is given by

$$h(50) = \frac{\Phi\left(\frac{\ln 50 - 5}{1}\right)}{50(1)(0.8621)} = 0.032 \text{ failures/unit.}$$

Thus, values for the log normal distribution are easily computed by using the standard normal tables.

Example 2.7: The failure time of a part is log normal distributed with μ = 6 and $\sigma =$ 2. Find the part reliability for a life of 200 time units.

Solution: The reliability for the part of 200 time units is

$$R(200) = P\left(Z > \frac{\ln 200 - 6}{2}\right) = P(Z > -0.35)$$
$$= 0.6368$$

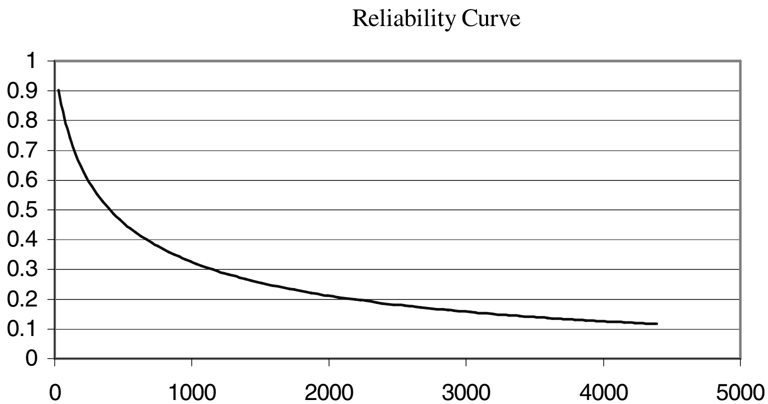

Figure 2.3. Log normal reliability plot *vs* time

The log normal lifetime model, like the normal, is flexible enough to make it a very useful empirical model. Figure 2.3 shows the reliability of the log normal vs time. It can be theoretically derived under assumptions matching many failure mechanisms. Some of these are: corrosion and crack growth, and in general, failures resulting from chemical reactions or processes.

Weibull Distribution

The exponential distribution is often limited in applicability owing to the memoryless property. The Weibull distribution (Weibull 1951) is a generalization of the exponential distribution and is commonly used to represent fatigue life, ball bearing life, and vacuum tube life. The Weibull distribution is extremely flexible and appropriate for modeling component lifetimes with fluctuating hazard rate functions and for representing various types of engineering applications. The three-parameters probability density function is

$$f(t) = \frac{\beta(t-\gamma)^{\beta-1}}{\theta^{\beta}} e^{-\left(\frac{t-\gamma}{\theta}\right)^{\beta}} \qquad t \geq \gamma \geq 0$$

where θ and β are known as the scale and shape parameters, respectively, and γ is known as the location parameter. These parameters are always positive. By using different parameters, this distribution can follow the exponential distribution, the normal distribution, *etc*. It is clear that, for $t \geq \gamma$, the reliability function $R(t)$ *is*

$$R(t) = e^{-\left(\frac{t-\gamma}{\theta}\right)^{\beta}} \quad \text{for } t > \gamma > 0, \beta > 0, \theta > 0 \tag{2.13}$$

hence,

$$h(t) = \frac{\beta(t-\gamma)^{\beta-1}}{\theta^{\beta}} \qquad t > \gamma > 0, \beta > 0, \theta > 0 \tag{2.14}$$

It can be shown that the hazard function is decreasing for $\beta < 1$, increasing for $\beta > 1$, and constant when $\beta = 1$.

Example 2.8: The failure time of a certain component has a Weibull distribution with $\beta = 4$, $\theta = 2000$, and $\gamma = 1000$. Find the reliability of the component and the hazard rate for an operating time of *1500* hours.

Solution: A direct substitution into equation (2.13) yields

$$R(1500) = e^{-\left(\frac{1500-1000}{2000}\right)^{4}} = 0.996$$

Using equation (2.14), the desired hazard function is given by

$$h(1500) = \frac{4(1500-1000)^{4-1}}{(2000)^{4}}$$

$$= 3.13 \text{ x } 10^{-5} \text{ failures/hour}$$

Note that the Rayleigh and exponential distributions are special cases of the Weibull distribution at $\beta = 2$, $\gamma = 0$, and $\beta = 1$, $\gamma = 0$, respectively. For example, when $\beta = 1$ and $\gamma = 0$, the reliability of the Weibull distribution function in equation (2.13) reduces to

$$R(t) = e^{-\frac{t}{\theta}}$$

and the hazard function given in equation (2.14) reduces to $1/\theta$, a constant. Thus, the exponential is a special case of the Weibull distribution. Similarly, when $\gamma = 0$ and $\beta = 2$, the Weibull probability density function becomes the Rayleigh density function. That is

$$f(t) = \frac{2}{\theta} t e^{-\frac{t^2}{\theta}} \quad \text{for } \theta > 0,\ t \geq 0$$

Other Forms of Weibull Distributions

The Weibull distribution again is widely used in engineering applications. It was originally proposed for representing the distribution of the breaking strength of materials. The Weibull model is very flexible and also has theoretical justification in many applications as a purely empirical model. Another form of Weibull probability density function is, for example,

$$f(x) = \lambda\ \gamma\ x^{\gamma-1}\ e^{-\lambda t^{\gamma}} \tag{2.15}$$

When $\gamma=2$, the density function becomes a Rayleigh distribution.

It can easily be shown that the mean, variance and reliability of the above Weibull distribution are, respectively, as follows:

$$\begin{aligned} \text{Mean} \quad &= \lambda^{-\frac{1}{\gamma}} \Gamma(1+\frac{1}{\gamma}) \\ \text{Variance} \quad &= \lambda^{-\frac{2}{\gamma}} \left(\Gamma\left(1+\frac{2}{\gamma}\right) - \left(\Gamma\left(1+\frac{1}{\gamma}\right)\right)^2 \right) \\ \text{Reliability} \quad &= e^{-\lambda t^{\gamma}} \end{aligned} \tag{2.16}$$

Example 2.9: The time to failure of a part has a Weibull distribution with $\frac{1}{\lambda}$ =250 (measured in 10^5 cycles) and $\gamma=2$. Find the part reliability at 10^6 cycles.

Solution: The part reliability at 10^6 cycles is

$$R(10^6) = e^{-(10)^2/250} = 0.6703$$

The resulting reliability function is shown in Figure 2.4.

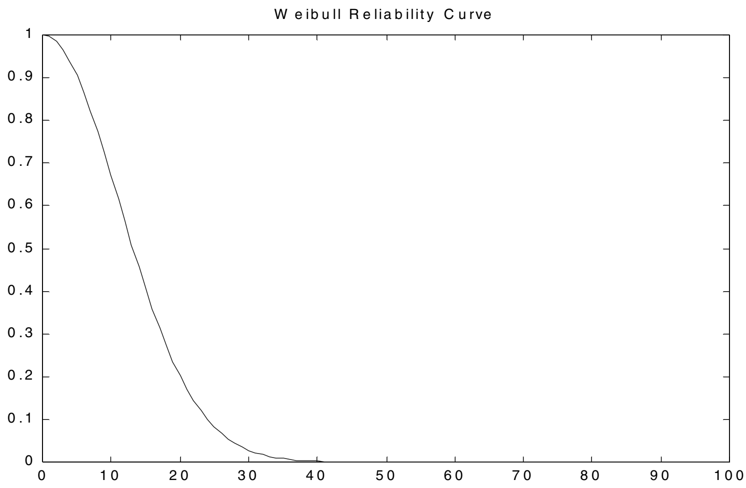

Figure 2.4. Weibull reliability function *vs* time

Gamma Distribution

Gamma distribution can be used as a failure probability function for components whose distribution is skewed. The failure density function for a gamma distribution is

$$f(t) = \frac{t^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-\frac{t}{\beta}} \qquad t \geq 0,\ \alpha,\ \beta > 0 \tag{2.17}$$

where α is the shape parameter and β is the scale parameter. Hence,

$$R(t) = \int_t^{\infty} \frac{1}{\beta^{\alpha}\ \Gamma(\alpha)} s^{\alpha-1}\ e^{-\frac{s}{\beta}} ds$$

If α is an integer, it can be shown by successive integration by parts that

$$R(t) = e^{-\frac{t}{\beta}} \sum_{i=0}^{\alpha-1} \frac{(\frac{t}{\beta})^i}{i!} \tag{2.18}$$

and

$$h(t) = \frac{f(t)}{R(t)} = \frac{\frac{1}{\beta^{\alpha}\ \Gamma(\alpha)}\ t^{\alpha-1}\ e^{-\frac{t}{\beta}}}{e^{-\frac{t}{\beta}} \sum_{i=0}^{\alpha-1} \frac{(\frac{t}{\beta})^i}{i!}}$$

The gamma density function has shapes that are very similar to the Weibull distribution. At α = 1, the gamma distribution becomes the exponential distribution with the constant failure rate $1/\beta$. The gamma distribution can also be used to model the time to the n^{th} failure of a system if the underlying failure distribution is exponential. Thus, if X_i is exponentially distributed with parameter $\theta = 1/\beta$, then $T = X_1 + X_2 + \ldots + X_n$, is gamma distributed with parameters β and n.

Example 2.10: The time to failure of a component has a gamma distribution with α = 3 and β = 5. Determine the reliability of the component and the hazard rate at 10 time-units.

Solution: Using equation (2.18), we compute

$$R(10) = e^{-\frac{10}{5}} \sum_{i=0}^{2} \frac{\left(\frac{10}{5}\right)^i}{i!} = 0.6767$$

From equation (2.17), we obtain

$$h(10) = \frac{f(10)}{R(10)} = \frac{0.054}{0.6767} = 0.798 \text{ failures/unit time}$$

The other form of the gamma probability density function can be written as follows:

$$f(x) = \frac{\beta^{\alpha}\, t^{\alpha-1}}{\Gamma(\alpha)} e^{-t\beta} \qquad \text{for } t>0 \tag{2.19}$$

This pdf is characterized by two parameters: shape parameter α and scale parameter β. When $0<\alpha<1$, the failure rate monotonically decreases; when $\alpha>1$, the failure rate monotonically increase; when $\alpha=1$ the failure rate is constant.

The mean, variance and reliability of the density function in equation (2.19) are, respectively,

$$\text{Mean(MTTF)} = \frac{\alpha}{\beta}$$

$$\text{Variance} = \frac{\alpha}{\beta^2}$$

$$\text{Reliability} = \int_t^{\infty} \frac{\beta^{\alpha}\, x^{\alpha-1}}{\Gamma(\alpha)} e^{-x\beta}\, dx$$

Example 2.11: A mechanical system time to failure is gamma distribution with $\alpha=3$ and $1/\beta=120$. Find the system reliability at 280 hours.

Solution: The system reliability at 280 hours is given by

$$R(280) = e^{\frac{-280}{120}} \sum_{k=0}^{2} \frac{\left(\frac{280}{120}\right)^2}{k!} = 0.85119$$

and the resulting reliability plot is shown in Figure 2.5.

The gamma model is a flexible lifetime model that may offer a good fit to some sets of failure data. It is not, however, widely used as a lifetime distribution model for common failure mechanisms. A common use of the gamma lifetime model occurs in Bayesian reliability applications.

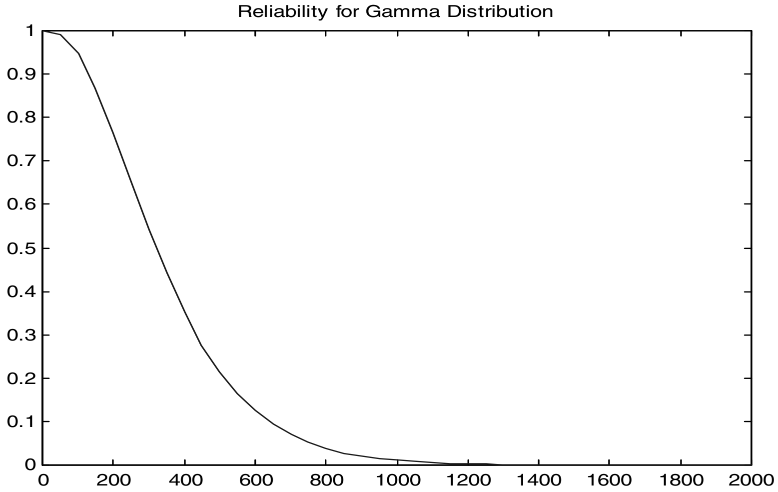

Figure 2.5. Gamma reliability function *vs* time

Beta Distribution

The two-parameter Beta density function, $f(t)$, is given by

$$f(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha}(1-t)^{\beta} \quad 0 < t < 1,\ \alpha > 0,\ \beta > 0$$

where α and β are the distribution parameters. This two-parameter distribution is commonly used in many reliability engineering applications.

Pareto Distribution

The Pareto distribution was originally developed to model income in a population. Phenomena such as city population size, stock price fluctuations, and personal incomes have distributions with very long right tails. The probability density function of the Pareto distribution is given by

$$f(t) = \frac{\alpha k^{\alpha}}{t^{\alpha+1}} \qquad k \le t \le \infty$$

The mean, variance and reliability of the Pareto distribution are, respectively,

Mean $= k/(\alpha-1)$ for $\alpha > 1$

Variance $= \dfrac{\alpha K^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$

Reliability $= \left(\dfrac{k}{t}\right)^{\alpha}$

The Pareto and log normal distributions have been commonly used to model the population size and economical incomes. The Pareto is used to fit the tail of the distribution, and the log normal is used to fit the rest of the distribution.

Rayleigh Distribution

The Rayleigh function is a flexible lifetime distribution that can apply to many degra- dation process failure modes. The Rayleigh probability density function is

$$f(t) = \frac{t}{\sigma^2} e^{\left(\frac{-t^2}{2\sigma^2}\right)} \tag{2.20}$$

The mean, variance, and reliability of Rayleigh function are, respectively,

$$\text{Mean} = \sigma\left(\frac{\pi}{2}\right)^{\frac{1}{2}}$$

$$\text{Variance} = \left(2-\frac{\pi}{2}\right)\sigma^2$$

$$\text{Reliability} = e^{\frac{-\sigma t^2}{2}}$$

Example 2.12: Rolling resistance is a measure of the energy lost by a tire under load when it resists the force opposing its direction of travel. In a typical car, traveling at 60 miles per hour, about 20% of the engine power is used to overcome the rolling resistance of the tires.

A tire manufacturer introduces a new material that, when added to the tire rubber compound, significantly improves the tire rolling resistance but increases the wear rate of the tire tread. Analysis of a laboratory test of 150 tires shows that the failure rate of the new tire linearly increases with time (hours). It is expressed as

$$h(t) = 0.5\times 10^{-8} t$$

Find the reliability of the tire at one year.

Solution: The reliability of the tire after one year (8760 hours) of use is

$$R(1_{year}) = e^{-\frac{0.5}{2}\times 10^{-8}\times (8760)^2} = 0.8254$$

Figure 2.6 shows the resulting reliability function.

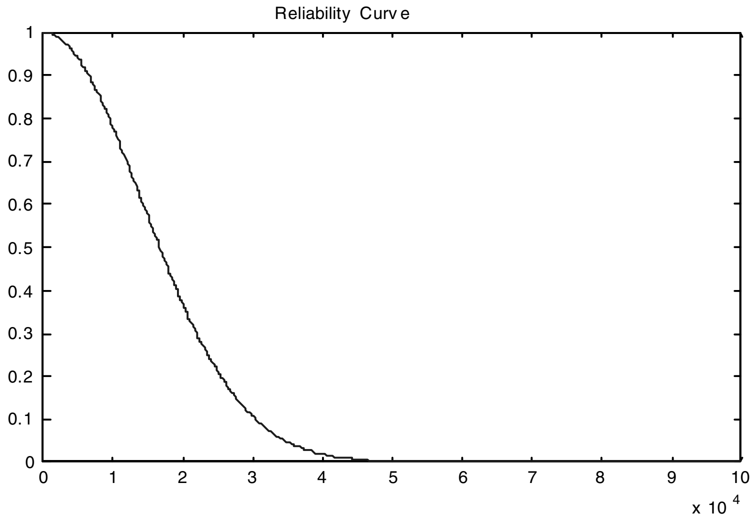

Figure 2.6. Rayleigh reliability function *vs* time

Vtub-shaped Hazard Rate Distribution

Pham (2002a) recently developed a two-parameter lifetime distribution with a Vtub-shaped hazard rate, called *Pham distribution* - also known as *Loglog distribution.*

Note that the loglog distribution with Vtub-shaped and Weibull distribution with bathtub-shaped failure rates are not the same. As for the bathtub-shaped, after the infant mortality period, the useful life of the system begins. During its useful life, the system fails as a constant rate. This period is then followed by a wear out period during which the system starts slowly and increases with the onset of wear out. For the Vtub-shaped, after the infant mortality period, the system starts to experience at a relatively low increasing rate, but this is not constant, and then increases with failures due to aging.

The Pham probability density function is given as follows (Pham 2002a):

$$f(t) = \alpha \ \ln a \ t^{\alpha-1} \ a^{t^\alpha} \ e^{1-a^{t^\alpha}} \quad \text{for } t>0,\ a>0,\ \alpha>0. \tag{2.21}$$

The Pham distribution and reliability functions are

$$F(t) = \int_0^t f(x)\, dx = 1 - e^{1-a^{t^\alpha}}$$

and

$$R(t) = e^{1-a^{t^\alpha}} \tag{2.22}$$

respectively. The corresponding failure rate of the Pham distribution is given by

$$h(t) = \alpha \ln(a)\ t^{\alpha-1}\ a^{t^\alpha} \tag{2.23}$$

Figures 2.7 and 2.8 describe the density function and failure rate function for various values of a and α.

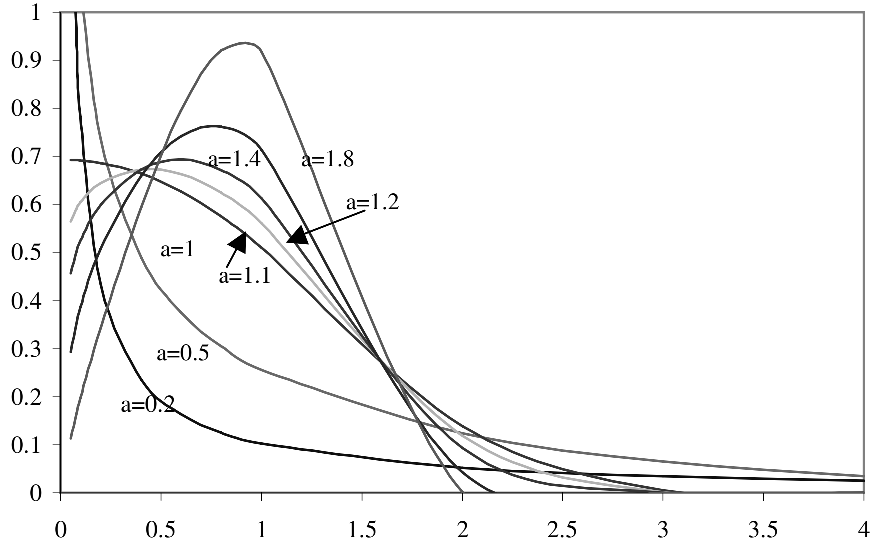

Figure 2.7. Probability density function for various values α with a=2

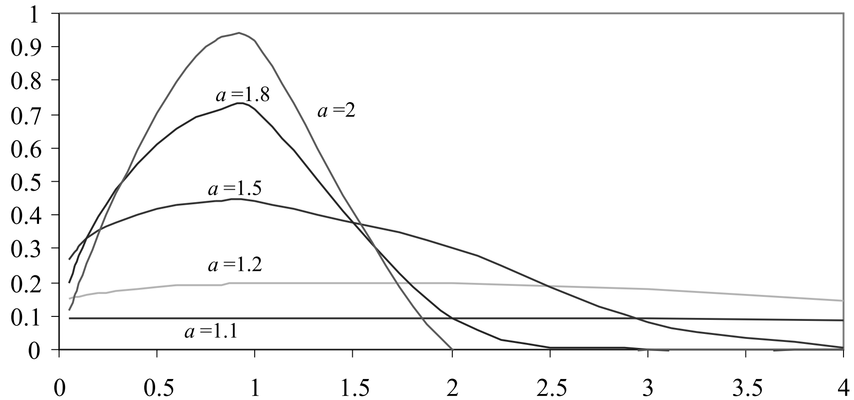

Figure 2.8. Probability density function for various values a with $\alpha = 1.5$

Two-Parameter Hazard Rate Function

This is a two-parameter function that can have increasing and decreasing hazard rates. The hazard rate, $h(t)$, the reliability function, $R(t)$, and the pdf are, respectively, given as follows

$$h(t) = \frac{n\lambda t^{n-1}}{\lambda t^n + 1} \quad \text{for } n \geq 1, \lambda > 0, t \geq 0 \tag{2.24}$$

$$R(t) = e^{-\ln(\lambda t^N + 1)} \tag{2.25}$$

and

$$f(t) = \frac{n\lambda t^{n-1}}{\lambda t^n + 1} e^{-\ln(\lambda t^n + 1)} \qquad n \geq 1, \lambda > 0, t \geq 0 \tag{2.26}$$

where n = shape parameter; λ = scale parameter

Three-Parameter Hazard Rate Function

This is a three-parameter distribution that can have increasing and decreasing hazard rates. The hazard rate, $h(t)$, is given as

$$h(t) = \frac{\lambda(b+1)[\ln(\lambda t + \alpha)]^b}{(\lambda t + \alpha)} \qquad b \geq 0, \lambda > 0, \alpha \geq 0, t \geq 0 \tag{2.27}$$

The reliability function $R(t)$ for $\alpha = 1$ is

$$R(t) = e^{-(\ln(\lambda t + \alpha))^{b+1}}$$

The probability density function $f(t)$ *is*

$$f(t) = e^{-[\ln(\lambda t + \alpha)]^{b+1}} \frac{\lambda(b+1)[\ln(\lambda t + \alpha)]^b}{(\lambda t + \alpha)} \tag{2.28}$$

where b = shape parameter, λ = scale parameter, and α = location parameter.

2.3 A Generalized Systemability Function

The traditional reliability definitions and its calculations have commonly been carried out through the failure rate function within a controlled laboratory-test environment. In other words, such reliability functions are applied to the failure testing data and then utilized to make predictions on the reliability of the system used in the field. The underlying assumption for such calculation is that the field environments and the testing environments are the same.

By defintion, a mathematical reliability function is the probability that a system will be successful in the interval from time 0 to time t, given by

$$R(t) = \int_t^\infty f(s)ds = e^{-\int_0^t h(s)ds} \tag{2.29}$$

where $f(s)$ and $h(s)$ are, respectively, the failure time density and failure rate function.

The operating environments are, however, often unknown and yet different due to the uncertainties of environments in the field (Pham and Xie 2003). A new look at how reliability researchers can take account of the randomness of the field environments into mathematical reliability modeling covering system failure in the field is great interest.

Pham (2005a) recently developed a new mathematical function called *systemability,* considering the uncertainty of the operational environments in the function for predicting the reliability of systems.

Notation

$h_i(t)$ i^{th} component hazard rate function

$R_i(t)$ i^{th} component reliability function

λ_i Intensity parameter of Weibull distribution for i^{th} component

$\underline{\lambda}$ $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3 ..., \lambda_n)$.

γ_i Shape parameter of Weibull distribution for i^{th} component

$\underline{\gamma}$ $\underline{\gamma} = (\gamma_1, \gamma_2, \gamma_3 ..., \gamma_n)$.

η A common environment factor

$G(\eta)$ Cumulative distribution function of η

α Shape parameter of Gamma distribution

β Scale parameter of Gamma distribution

2.3.1 Systemability Definition

This section discusses a definition of systemability function.

Definition 2.2 (Pham 2005a): *Systemability* is defined as the probability that the system will perform its intended function for a specified mission time under the random operational environments.

In a mathematical form, the *systemabililty* function is given by

$$R_s(t) = \int_{\eta} e^{-\eta \int_0^t h(s)ds} dG(\eta) \tag{2.30}$$

where η is a random variable that represents the system operational environments with a distribution function *G*.

This new function captures the uncertainty of complex operational environments of systems in terms of the system failure rate. It also would reflect the reliability estimation of the system in the field.

If we assume that η has a gamma distribution with parameters α and β , *i.e.*, $\eta \sim gamma(\alpha, \beta)$ where the pdf of η is given by

$$f_\eta(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad \text{for } \alpha,\ \beta>0;\ x \geq 0 \tag{2.31}$$

then the systemability function of the system in equation (2.30), using the Laplace transform (see Appendix 2), is given by

$$R_s(t) = \left[\frac{\beta}{\beta + \int_0^t h(s)ds} \right]^\alpha \tag{2.32}$$

2.3.2 Systemability Calculations

This subsection presents several systemability results and variances of some system configurations such as series, parallel, and *k*-out-of-*n* systems (Pham 2005a). Consider the following assumptions:

1. A system consists of *n* independent components where the system is subject to a random operational environment η.
2. i^{th} component lifetime is assumed to follow the Weibull density function, *i.e.*

Component hazard rate $h_i(t) = \lambda_i \; \gamma_i \; t^{\gamma_i - 1}$ (2.33)

Component reliability $R_i(t) = e^{-\lambda_i \, t^{\gamma_i}} \qquad t > 0$ (2.34)

Given common environment factor $\eta \sim gamma(\alpha, \beta)$, the systemability functions for different system structures can be obtained as follows.

Series System Configuration

In a series system, all components must operate successfully if the system is to function. The conditional reliability function of series systems subject to an actual operational random environment η is given by

$$R_{Series}(t \mid \eta, \underline{\lambda}, \underline{\gamma}) = e^{\left(-\eta \sum_{i=1}^{n} \lambda_i \, t^{\gamma_i}\right)} \tag{2.35}$$

The series systemability is given as follows

$$R_{Series}(t \mid \underline{\lambda}, \underline{\gamma}) = \int_{\eta} \exp\left(-\eta \sum_{i=1}^{n} \lambda_i t^{\gamma_i}\right) dG(\eta) = \left[\frac{\beta}{\beta + \sum_{i=1}^{n} \lambda_i t^{\gamma_i}}\right]^{\alpha} \tag{2.36}$$

The variance of a general function *R*(*t*) is given by

$$Var\left[R(t)\right] = E\left[R^2(t)\right] - \left(E\left[R(t)\right]\right)^2 \tag{2.37}$$

Given $\eta \sim gamma(\alpha, \beta)$, the variance of systemability for any system structure can be easily obtained. Therefore, the variance of series systemability is given by

$$\begin{aligned} Var\left[R_{Series}(t \mid \underline{\lambda}, \underline{\gamma})\right] = \int_{\eta} \exp\left(-\eta\left(2\sum_{i=1}^{n} \lambda_i t^{\gamma_i}\right)\right) dG(\eta) - \\ \left(\int_{\eta} \exp\left(-\eta \sum_{i=1}^{n} \lambda_i t^{\gamma_i}\right) dG(\eta)\right)^2 \end{aligned} \tag{2.38}$$

or

$$Var\left[R_{Series}(t \mid \underline{\lambda}, \underline{\gamma})\right] = \left[\frac{\beta}{\beta + 2\sum_{i=1}^{n} \lambda_i t^{\gamma_i}}\right]^{\alpha} - \left[\frac{\beta}{\beta + \sum_{i=1}^{n} \lambda_i t^{\gamma_i}}\right]^{2\alpha} \tag{2.39}$$

Parallel System Configuration

A parallel system is a system that is not considered to have failed unless all components have failed. The conditional reliability function of parallel systems subject to the uncertainty operational environment η is given by

$$R_{Parallel}(t\mid\eta,\underline{\lambda},\underline{\gamma}) = \exp\left(-\eta\lambda_i t^{\gamma_i}\right) - \sum_{\substack{i_1,i_2=1\\ i_1\neq i_2}}^{n} \exp\left(-\eta\left(\lambda_{i_1} t^{\gamma_{i_1}} + \lambda_{i_2} t^{\gamma_{i_2}}\right)\right) +$$
$$\sum_{\substack{i_1,i_2,i_2=1\\ i_1\neq i_2\neq i_3}}^{n} \exp\left(-\eta\left(\lambda_{i_1} t^{\gamma_{i_1}} + \lambda_{i_2} t^{\gamma_{i_2}} + \lambda_{i_3} t^{\gamma_{i3}}\right)\right) - \tag{2.40}$$
$$\cdots\cdots$$
$$+(-1)^{n-1}\exp\left(-\eta\sum_{i=1}^{n}\lambda_i t^{\gamma_i}\right)$$

Hence, the parallel systemability is given by

$$R_{parallel}(t\mid\underline{\lambda},\underline{\gamma}) = \sum_{i=1}^{n}\left[\frac{\beta}{\beta+\lambda_i t^{\gamma_i}}\right]^{\alpha} - \sum_{\substack{i_1,i_2=1\\ i_1\neq i_2}}^{n}\left[\frac{\beta}{\beta+\lambda_{i_1} t^{\gamma_{i_1}} + \lambda_{i_2} t^{\gamma_{i_2}}}\right]^{\alpha} +$$
$$\sum_{\substack{i_1,i_2,i_2=1\\ i_1\neq i_2\neq i_3}}^{n}\left[\frac{\beta}{\beta+\lambda_{i_1} t^{\gamma_{i_1}} + \lambda_{i_2} t^{\gamma_{i_2}} + \lambda_{i_3} t^{\gamma_{i3}}}\right]^{\alpha} - \tag{2.41}$$
$$\cdots\cdots$$
$$+(-1)^{n-1}\left[\frac{\beta}{\beta+\sum_{i=1}^{n}\lambda_i t^{\gamma_i}}\right]^{\alpha}$$

or

$$\mathrm{R}_{parallel}(t\mid\underline{\lambda},\underline{\gamma}) = \sum_{k=1}^{n}(-1)^{k-1}\sum_{\substack{i_1,i_2\ldots,i_k=1\\ i_1\neq i_2\ldots\neq i_k}}^{n}\left[\frac{\beta}{\beta+\sum_{j=i_1,\ldots i_k}\lambda_j t^{\gamma_j}}\right]^{\alpha} \tag{2.42}$$

To simplify the calculation of a general n-component parallel system, we only consider here a parallel system consisting of two components. It is easy to see that the second-order moments of the systemability function can be written as

$$E\left[\mathrm{R}^{2}_{\mathrm{Parallel}}(t\mid\underline{\lambda},\underline{\gamma})\right] = \int_{\eta}(e^{-2\eta\,\lambda_1 t^{\gamma 1}} + e^{-2\eta\,\lambda_2\,t^{\gamma 2}} + \mathrm{e}^{-2\eta(\lambda_1 t^{\gamma 1}+\lambda_2 t^{\gamma 2})}$$
$$+\,\mathrm{e}^{-\eta(\lambda_1 t^{\gamma 1}+\lambda_2 t^{\gamma 2})} - e^{-\eta(2\lambda_1 t^{\gamma 1}+\lambda_2 t^{\gamma 2})} - \mathrm{e}^{-\eta(\lambda_1 t^{\gamma 1}+2\lambda_2 t^{\gamma 2})})\, d\ G(\eta)$$

The variance of series systemability of a two-component parallel system is given by

$$
\begin{aligned}
Var\left[R_{Parallel}(t\,|\,\underline{\lambda},\underline{\gamma})\right] = & \left[\frac{\beta}{\beta+2\lambda_1 t^{\gamma_1}}\right]^{\alpha} + \left[\frac{\beta}{\beta+2\lambda_2 t^{\gamma_2}}\right]^{\alpha} + \\
& \left[\frac{\beta}{\beta+2\lambda_1 t^{\gamma_1}+2\lambda_2 t^{\gamma_2}}\right]^{\alpha} + \left[\frac{\beta}{\beta+\lambda_1 t^{\gamma_1}+\lambda_2 t^{\gamma_2}}\right]^{\alpha} - \\
& \left[\frac{\beta}{\beta+2\lambda_1 t^{\gamma_1}+\lambda_2 t^{\gamma_2}}\right]^{\alpha} - \left[\frac{\beta}{\beta+\lambda_1 t^{\gamma_1}+2\lambda_2 t^{\gamma_2}}\right]^{\alpha} - \\
& \left[\left[\frac{\beta}{\beta+\lambda_1 t^{\gamma_1}}\right]^{\alpha} + \left[\frac{\beta}{\beta+\lambda_2 t^{\gamma_2}}\right]^{\alpha} - \left[\frac{\beta}{\beta+\lambda_1 t^{\gamma_1}+\lambda_2 t^{\gamma_2}}\right]^{\alpha}\right]^2
\end{aligned}
\tag{2.43}
$$

k-out-of-*n* System Configuration

In a *k*-out-of-*n* configuration, the system will operate if at least k out of n components are operating. To simplify the complexity of the systemability function, we assume that all the components in the *k*-out-of-*n* systems are identical. Therefore, for a given common environment η, the conditional reliability function of a component is given by

$$R(t\,|\,\eta,\lambda,\gamma) = e^{-\eta\lambda t^{\gamma}} \tag{2.44}$$

The conditional reliability function of *k*-out-of-*n* systems subject to the uncertainty operational environment η can be obtained as follows:

$$R_{k-out-of-n}(t\,|\,\eta,\lambda,\gamma) = \sum_{j=k}^{n}\binom{n}{j} e^{-\eta j\lambda t^{\gamma}} (1-e^{-\eta\lambda t^{\gamma}})^{(n-j)} \tag{2.45}$$

Note that

$$(1-e^{-\eta\lambda t^{\gamma}})^{(n-j)} = \sum_{l=0}^{n-j}\binom{n-j}{l}(-e^{-\eta\lambda t^{\gamma}})^{l}$$

The conditional reliability function of *k*-out-of-*n* systems, from equation (2.45), can be rewritten as

$$R_{k-out-of-n}(t\,|\,\eta,\lambda,\gamma) = \sum_{j=k}^{n}\binom{n}{j}\sum_{l=0}^{n-j}\binom{n-j}{l}(-1)^{l} e^{-\eta(j+l)\lambda t^{\gamma}} \tag{2.46}$$

Then if $\eta \sim gamma(\alpha,\beta)$ then the *k*-out-of-*n* systemability is given by

$$R_{(T_1,\ldots,T_n)}(t\,|\,\lambda,\gamma) = \sum_{j=k}^{n}\binom{n}{j}\sum_{l=0}^{n-j}\binom{n-j}{l}(-1)^{l}\left[\frac{\beta}{\beta+\lambda(j+l)t^{\gamma}}\right]^{\alpha} \tag{2.47}$$

It can be easily shown that

$$R^2_{k-out-of-n}(t\,|\,\eta,\lambda,\gamma) = \sum_{i=k}^{n}\binom{n}{i}\sum_{j=k}^{n}\binom{n}{j} e^{-\eta(i+j)\lambda t^{\gamma}} (1-e^{-\eta\lambda t^{\gamma}})^{(2n-i-j)} \tag{2.48}$$

Since

$$(1-e^{-\eta\lambda t^{\gamma}})^{(2n-i-j)} = \sum_{l=0}^{2n-i-j}\binom{2n-i-j}{l}(-e^{-\eta\lambda t^{\gamma}})^{l} \tag{2.49}$$

we can rewrite equation (2.48), after several simplifications, as follows

$$R^2_{k-out-of-n}(t\,|\,\eta,\lambda,\gamma) = \sum_{i=k}^{n}\binom{n}{i}\sum_{j=k}^{n}\binom{n}{j}(-1)^{l}\sum_{l=0}^{2n-i-j}\binom{2n-i-j}{l}e^{-\eta(i+j+l)\lambda t^{\gamma}} \tag{2.50}$$

Therefore, the variance of k-out-of-n system systemability function is given by

$$Var(R_{k/n}(t\,|\,\lambda,\gamma) = \int_{\eta} R^2_{k/n}(t\,|\,\eta,\lambda,\gamma)dG(\eta) - \left[\int_{\eta} R_{k/n}(t\,|\,\eta,\lambda,\gamma)dG(\eta)\right]^2$$

$$=\sum_{i=k}^{n}\binom{n}{i}\sum_{j=k}^{n}\binom{n}{j}\sum_{l=0}^{2n-i-j}\binom{2n-i-j}{l}(-1)^{l}\left(\frac{\beta}{\beta+(i+j+l)\lambda t^{\gamma}}\right)^{2} \tag{2.51}$$

$$-\left(\sum_{j=k}^{n}\binom{n}{j}\sum_{l=0}^{n-j}\binom{n-j}{l}(-1)^{l}\left(\frac{\beta}{\beta+(j+l)\lambda t^{\gamma}}\right)^{2}\right)^{2}$$

Example 2.13: Consider a k-out-of-n system where λ = 0.0001, $\gamma = 1.5$, n = 5, and $\eta \sim gamma(\alpha,\beta)$. Calculate the systemability of various k-out-of-n system configurations.

Solution: The systemability of generalized k-out-of-*5* system configurations is given as follows:

$$R_{k-out-of-n}(t\,|\,\lambda,\gamma) = \sum_{j=k}^{5}\binom{5}{j}\sum_{l=0}^{5-j}\binom{5-j}{l}(-1)^{l}\left[\frac{\beta}{\beta+\lambda(j+l)t^{\gamma}}\right]^{\alpha} \tag{2.52}$$

Figures 2.9 and 2.10 show the reliability function (conventional reliability function) and systemability function (equation 2.52) of a series system (here k=5) for $\alpha = 2, \beta = 3$ and for $\alpha = 2, \beta = 1$, respectively.

Figures 2.11 and 2.12 show the reliability and systemability functions of a parallel system (here k=1) for $\alpha = 2, \beta = 3$ and for $\alpha = 2, \beta = 1$, respectively. Similarly, Figures 2.13 and 2.14 show the reliability and systemability functions of a *3*-out-of-*5* system for $\alpha = 2, \beta = 3$ and for $\alpha = 2, \beta = 1$, respectively.

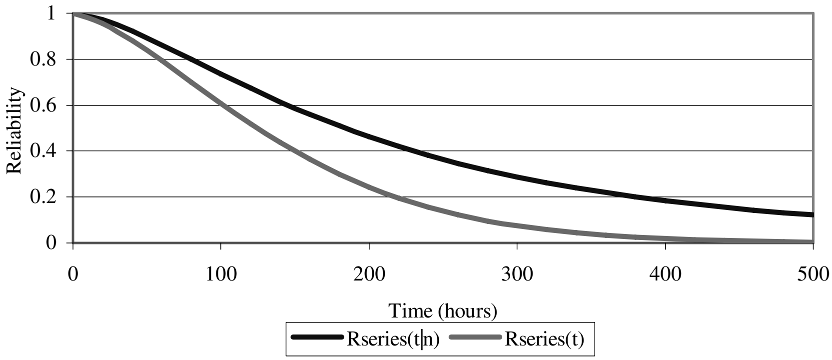

Figure 2.9. Comparisons of series system reliability vs systemability functions for $\alpha = 2$ and $\beta = 3$

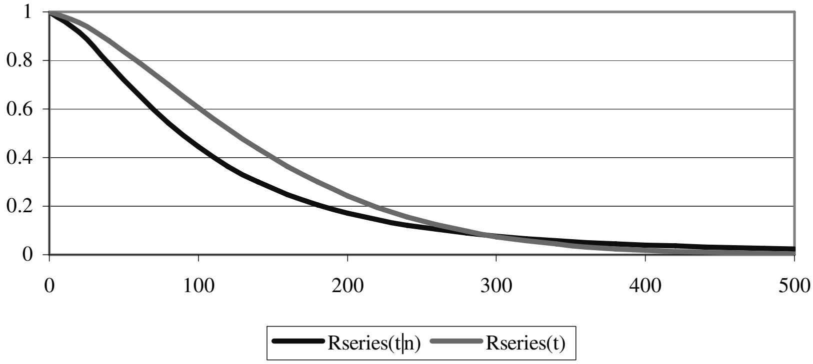

Figure 2.10. Comparisons of series system reliability vs. systemability functions for $\alpha = 2$ and $\beta = 1$

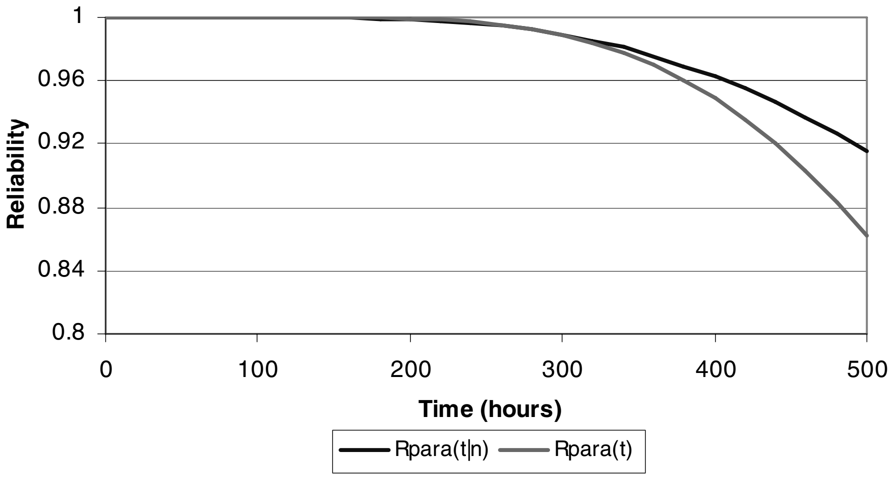

Figure 2.11. Comparisons of parallel system reliability *vs* systemability function for $\alpha = 2$ and $\beta = 3$

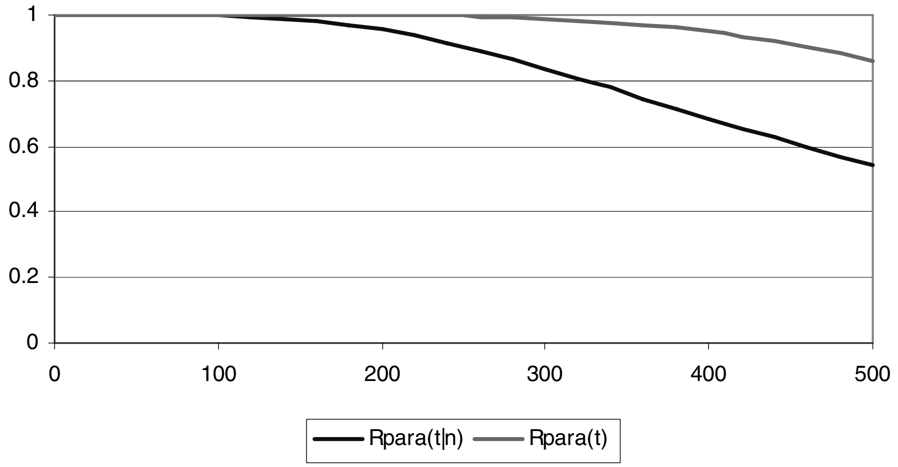

Figure 2.12. Comparisons of parallel system reliability *vs* Systemability functions for $\alpha = 2$ and $\beta = 1$

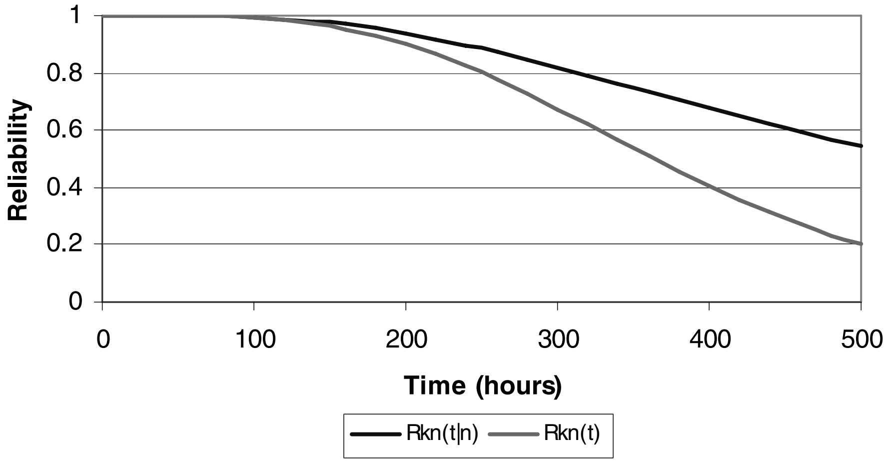

Figure 2.13. Comparisons of k-out-of-n system reliability vs. systemability functions for $\alpha = 2$ and $\beta = 3$

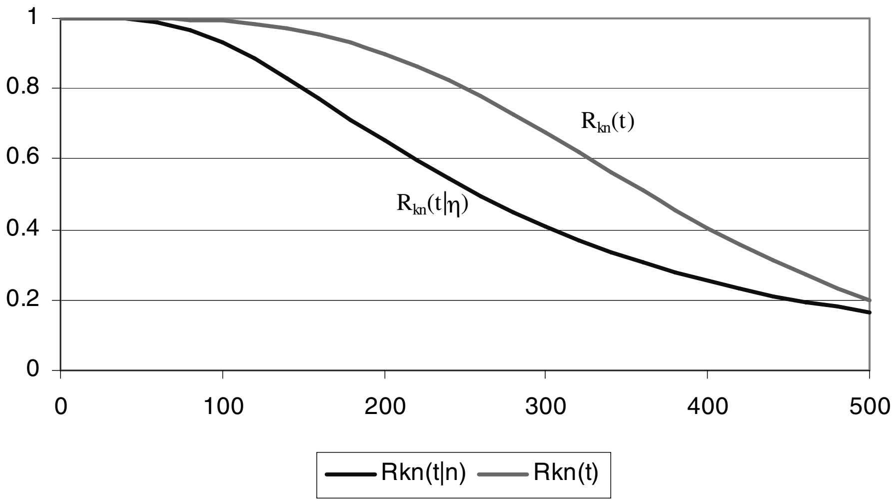

Figure 2.14. Comparisons of k-out-of-n system reliability vs. systemability functions for $\alpha = 2$ and $\beta = 1$

Variance of Systemability Calculations

Assume $\lambda = 0.00001$, $\gamma_{.} = 1.5$, n = 3, k = 2, and $\eta \sim gamma(\alpha, \beta)$, Figures 2.15 and 2.16 shows the systemability and its confidence intervals of a *2*-out-of-*3* system (Pham 1993) for $\alpha = 2, \beta = 1$ and $\alpha = 2, \beta = 2$, respectively.

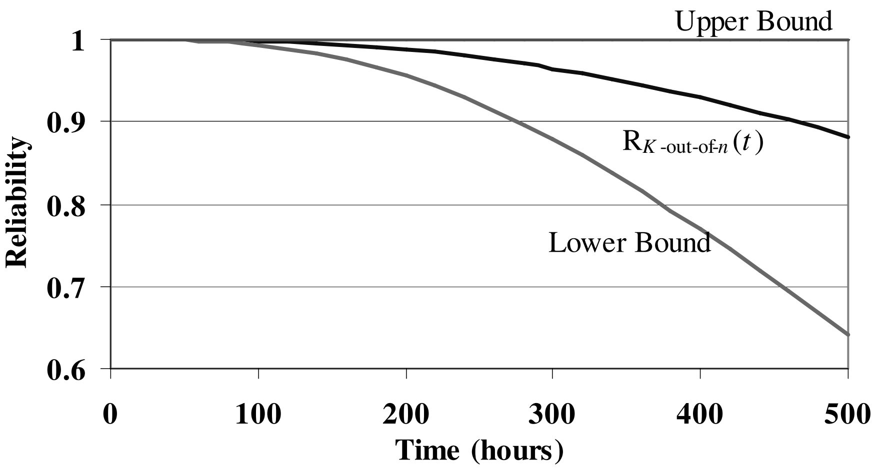

Figure 2.15. A *2*-out-of-*3* systemability and its 95% confidence interval where $\alpha = 2$, $\beta = 1$

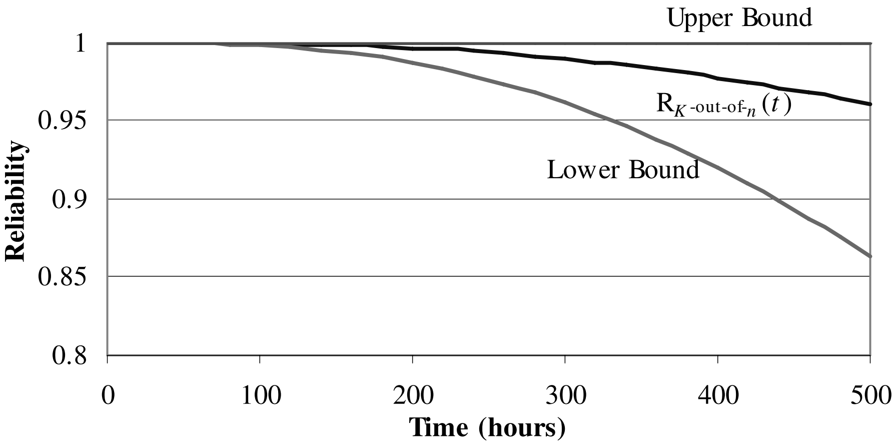

Figure 2.16. A *2*-out-of-*3* systemability and its 95% confidence interval ($\alpha = 2$, $\beta = 2$)

2.4 System Reliability with Multiple Failure Modes

This section discusses various reliability and optimization aspects of systems subject to multiple types of failure. It is assumed that the system component states are statistically independent and identically distributed. Networks of relays, diode circuits, fluid flow valves, *etc.* are a few examples of systems having components subject to failure in either open or closed modes.

The designations "closed mode" and "short mode" both appear in this section, and we will use the two terms interchangeably. Redundancy can be used to enhance the reliability of a system without any change in the reliability of the individual components that form the system. However, in a two-failure mode problem, redundancy may either increase or decrease the system's reliability. Therefore, adding components to the system may not increase the system reliability.

The reliability of a system subject to two kinds of failure is calculated as follows (Malon 1989):

System reliability = Pr{system works in both modes} (2.53)
= Pr{system works in open mode} - Pr{system fails in closed mode}+ Pr{system fails in both modes)

When the open- and closed-mode failure structures are dual of one another, *i.e.* Pr{system fails in both modes} = 0, then the system reliability given by equation (2.53) becomes

System reliability = 1 - Pr{system fails in open mode}
- Pr{system fails in closed mode} (2.54)

Notation

q_0	The open-mode failure probability of each component ($p_0 = 1 - q_0$)
q_s	The short-mode failure probability of each component ($p_s = 1 - q_s$)
$\lfloor x \rfloor$	The largest integer not exceeding x
*	Implies an optimal value

2.4.1 Reliability Calculations

The Series System

Consider a series system consisting of n components. In this series system, any one component failing in an open mode causes system failure in open mode whereas all components of the system must malfunction in short mode for the system to fail in closed mode.

The probabilities of system fails in open mode and fails in short mode are

$$F_0(n) = 1 - (1 - q_0)^n$$

and

$$F_s(n) = q_s^n$$

respectively. From equation (2.54), the system reliability is

$$R_s(n) = (1 - q_0)^n - q_s^n \tag{2.55}$$

where n is the number of identical and independent components. In a series arrangement, reliability with respect to closed system failure increases with the number of components, whereas reliability with respect to open system failure decreases.

Theorem 2.1: Let q_0 and q_s be fixed. There exists an optimum number of components, say n^*, that maximizes the system reliability. If we define

$$n_0 = \frac{\log\left(\dfrac{q_0}{1-q_s}\right)}{\log\left(\dfrac{q_s}{1-q_0}\right)}$$

then the system reliability, $R_s(n^*)$, is maximum for

$$n^* = \begin{cases} \lfloor n_0 \rfloor + 1 & \text{if } n_0 \text{ is not an integer} \\ n_0 \text{ or } n_0 + 1 & \text{if } n_0 \text{ is an integer} \end{cases} \tag{2.56}$$

Proof: The proof is left as an exercise for the reader (see Problem 2.17).

Example 2.14: A switch has two failure modes: fail-open and fail-short. The probability of switch open-circuit failure and short-circuit failure are 0.1 and 0.2 respectively. A system consists of n switches wired in series. That is, given q_0 = 0.1 and q_s = 0.2. Then

$$n_0 = \frac{\log\left(\dfrac{0.1}{1-0.2}\right)}{\log\left(\dfrac{0.2}{1-0.1}\right)} = 1.4$$

Thus, $n^* = \lfloor 1.4 \rfloor + 1 = 2$. Therefore, when $n^* = 2$ the system reliability $R_s(n)$ = 0.77 is maximized.

The Parallel System

Consider a parallel system consisting of n components. For a parallel configuration, all the components must fail in open mode or at least one component must malfunction in short mode to cause the system to fail completely. The system reliability is given by

$$R_p(n) = (1-q_s)^n - q_0^n \tag{2.57}$$

where n *is* the number of components connected in parallel. In this case, $(1 - q_s)^n$ represents the probability that no components fail in short mode, and q_0^n represents the probability that all components fail in open mode.

Theorem 2.2: If we define

$$n_0 = \frac{\log\left(\dfrac{q_s}{1-q_0}\right)}{\log\left(\dfrac{q_0}{1-q_s}\right)} \tag{2.58}$$

then the system reliability $R_p(n^*)$ is maximum for

$$n^* = \begin{cases} \lfloor n_0 \rfloor + 1 & \text{if } n_0 \text{ is not an integer} \\ n_0 \text{ or } n_0 + 1 & \text{if } n_0 \text{ is an integer.} \end{cases} \tag{2.59}$$

Proof: The proof is left as an exercise for the reader (see Problem 2.18).

It is observed that, for any range of q_0 and q_s, the optimal number of parallel components that maximizes the system reliability is one, if $q_s > q_0$ (see Problem 2.19). For most other practical values of q_0 and q_s the optimal number turns out to be two. In general, the optimal value of parallel components can be easily obtained using equation (2.58).

The Parallel-Series System

Consider a system of components arranged so that there are m subsystems operating in parallel, each subsystem consisting of n identical components in series. Such an arrangement is called a parallel-series arrangement. The components are subject to two types of failure: failure in open mode and failure in short mode.

The systems are characterized by the following properties:
1. The system consists of m subsystems, each subsystem containing n i.i.d. components.
2. A component is either good, failed open, or failed short. Failed components can never become good, and there are no transitions between the open and short failure modes.
3. The system can be (a) good, (b) failed open (at least one component in each subsystem fails open), or (c) failed short (all the components in any subsystem fail short).
4. The unconditional probabilities of component failure in open and short modes are known and are constrained: $q_o, q_s > 0$; $q_o + q_s < 1$.

The probabilities of a system failing in open mode and failing in short mode are given by

$$F_0(m) = [1-(1-q_0)^n]^m \tag{2.60}$$

and

$$F_s(m) = 1-(1-(q_s)^n)^m \tag{2.61}$$

respectively. The system reliability is

$$R_{ps}(n,m) = (1-q_s^n)^m - [1-(1-q_0)^n]^m \tag{2.62}$$

An interesting example in Barlow and Proschan (1965) shows that there exists no pair n, m maximizing system reliability, since R_{ps} is made arbitrarily close to one by appropriate choice of m and n. To see this, let

$$a = \frac{\log q_s - \log(1-q_0)}{\log q_s + \log(1-q_0)} \qquad M_n = q_s^{-n/(1+a)} \qquad m_n = \lfloor M_n \rfloor$$

For given n, take $m = m_n$; then one can rewrite equation (2.62) as:

$$R_{ps}(n, m_n) = (1 - q_s^n)^{m_n} - [1 - (1 - q_0)^n]^{m_n}$$

A straightforward computation yields

$$\lim_{n\to\infty} R_{ps}(n, m_n) = \lim_{n\to\infty}\{(1 - q_s^n)^{m_n} - [1 - (1 - q_0)^n]^{m_n}\} = 1$$

For fixed n, q_0, and q_s, one can determine the value of m that maximizes R_{ps}.

Theorem 2.3 (Barlow and Proschan 1965): Let n, q_0, and q_s be fixed. The maximum value of $R_{ps}(m)$ is attained at $m^* = \lfloor m_0 \rfloor + 1$, where

$$m_0 = \frac{n(\log p_0 - \log q_s)}{\log(1 - q_s^n + \log(1 - p_0^n)} \tag{2.63}$$

If m_o is an integer, then m_o and m_o+1 both maximize $R_{ps}(m)$.

Proof: The proof is left as an exercise for the reader (see Problem 20).

The Series-Parallel System

The series-parallel structure is the dual of the parallel-series structure. We consider a system of components arranged so that there are m subsystems operating in series, each subsystem consisting of n identical components in parallel. Such an arrangement is called a series-parallel arrangement.

Failure in open mode of all the components in any subsystem makes the system unresponsive. Failure in closed (short) mode of a single component in each subsystem also makes the system unresponsive. The probabilities of system failure in open and short mode are given by

$$F_0(m) = 1 - (1 - q_0^n)^m \tag{2.64}$$

and

$$F_s(m) = [1 - (1 - q_s)^n]^m \tag{2.65}$$

respectively. The system reliability is

$$R(m) = (1 - q_0^n)^m - [1 - (1 - q_s)^n]^m \tag{2.66}$$

where m is the number of identical subsystems in series and n is the number of identical components in each parallel subsystem.

Barlow and Proschan (1965) show that there exists no pair (m, n) maximizing system reliability. For fixed m, q_0, and q_s however, one can determine the value of n that maximizes the system reliability.

Theorem 2.4 (Barlow and Proschan 1965): Let n, q_0, and q_s be fixed. The maximum value of $R(m)$ is attained at $m^* = \lfloor m_0 \rfloor + 1$, where

$$m_0 = \frac{n(\log p_s - \log q_0)}{\log(1 - q_0^n) - \log(1 - p_s^n)} \tag{2.67}$$

If m_o is an integer, then m_o and $m_o + 1$ both maximize $R(m)$.

Proof: (see Problem 21).

The *k*-out-of-*n* Systems

Consider a k-out-of-n system consisting of n identical and independent components that can be either good or failed. The components are subject to two types of failure: failure in open mode and failure in closed mode. The k out of n system can fail when k or more components fail in closed mode or when $(n - k + 1)$ or more components fail in open mode.

Applications of k-out-of-n systems can be found in the areas of target detection, communication, and safety monitoring systems, and, particularly, in the area of human organizations. The following is an example in the area of human organizations (Nordmann and Pham 1999).

Consider a committee with n members who must decide to accept or reject innovation-oriented projects. The projects are of two types: "good" and "bad". It is assumed that the communication among the members is limited, and each member will make a yes-no decision on each project. A committee member can make two types of error: the error of accepting a bad project and the error of rejecting a good project. The committee will accept a project when k or more members accept it, and will reject a project when $(n - k + 1)$ or more members reject it.

Thus, the two types of potential error of the committee are: (1) the acceptance of a bad project (which occurs when k or more members make the error of accepting a bad project); (2) the rejection of a good project (which occurs when $(n - k + 1)$ or more members make the error of rejecting a good project).

This section determines the optimal k or n that maximizes the system reliability. We also study the effect of the system's parameters on the optimal k or n. The system fails in closed mode if and only if at least k of its n components fail in closed mode, and we obtain

$$F_s(k,n) = \sum_{i=k}^{n} \binom{n}{i} q_s^i p_s^{n-i} = 1 - \sum_{i=0}^{k-1} \binom{n}{i} q_s^i p_s^{n-i} \tag{2.68}$$

The system fails in open mode if and only if at least $(n - k + 1)$ of its n components fail in open mode, that is:

$$F_0(k,n) = \sum_{i=n-k+1}^{n} \binom{n}{i} q_0^i p_0^{n-i} = \sum_{i=0}^{k-1} \binom{n}{i} p_0^i q_0^{n-i} \tag{2.69}$$

The system reliability is given by

$$R(k,n) = 1 - F_0(k,n) - F_s(k,n) = \sum_{i=0}^{k-1} \binom{n}{i} q_s^i p_s^{n-i} - \sum_{i=0}^{k-1} \binom{n}{i} p_0^i q_0^{n-i} \tag{2.70}$$

For a given k, we can find the optimum value of n, say n^*, that maximizes the system reliability.

Theorem 2.5 (Pham 1989a): For fixed k, q_0, and q_s, the maximum value of R(k, n) is attained at $n^* = \lfloor n_0 \rfloor$ where

$$n_0 = k\left[1 + \frac{\log\left(\dfrac{1-q_0}{q_s}\right)}{\log\left(\dfrac{1-q_s}{q_0}\right)}\right] \tag{2.71}$$

If n_0 is an integer, both n_0 and $n_0 + 1$ maximize $R(k, n)$.

Proof: The proof is left as an exercise for the reader (see Problem 22).

This result shows that when n_0 is an integer, both n^*-1 and n^* maximize the system reliability $R(k, n)$. In such cases, the lower value will provide the more economical optimal configuration for the system. If $q_0 = q_s$ the system reliability $R(k, n)$ is maximized when $n = 2k$ or $2k$-1. In this case, the optimum value of n does not depend on the value of q_0 and q_s and the best choice for a decision voter is a majority voter; this system is also called a majority system (Pham,1989a).

From Theorem 2.5, we understand that the optimal system size n* depends on the various parameters q_0 and q_s. It can be shown the optimal value n* is an increasing function of q_0 and a decreasing function of q_s (see Problem 23). Intuitively, these results state that when q_s increases it is desirable to reduce the number of components in the system as close to the value of threshold level k as possible. On the other hand, when q_0 increases, the system reliability will be improved if the number of components increases.

Theorem 2.6 (Ben-Dov 1980): For fixed n, q_0, and q_s, it is straightforward to see that the maximum value of R(k, n) is attained at $k^* = \lfloor k_0 \rfloor + 1$, where

$$k_0 = n\frac{\log\left(\dfrac{q_0}{p_s}\right)}{\log\left(\dfrac{q_s q_0}{p_s p_0}\right)} \tag{2.72}$$

If k_0 is an integer, both k_0 and $k_0 + 1$ maximize R(k, n).

Proof: The proof is left as an exercise for the reader (see Problem 24).

We now discuss how these two values, k^* and n^*, are related to one another. Define α by

$$\alpha = \frac{\log\left(\dfrac{q_0}{p_s}\right)}{\log\left(\dfrac{q_s q_0}{p_s p_0}\right)} \tag{2.73}$$

then, for a given n, the optimal threshold k is given by $k^* = \lceil n\alpha \rceil$ and for a given k the optimal n *is* $n^* = \lfloor k/\alpha \rfloor$. For any given q_0 and q_s, we can easily show that (see Problem 25)

$$q_s < \alpha < p_0 \tag{2.74}$$

Therefore, we can obtain the following bounds for the optimal value of the threshold k:

$$nq_s < k^* < np_0 \tag{2.75}$$

This result shows that for given values of q_0 and q_s, an upper bound for the optimal threshold k^* is the expected number of components working in open mode, and a lower bound for the optimal threshold k^* is the expected number of components failing in closed mode.

2.4.2 An Application of Systems with Multiple Failure Modes

In many critical applications of digital systems, fault tolerance has been an essential architectural attribute for achieving high reliability. Several techniques can achieve fault tolerance using redundant hardware (Mathur and De Sousa 1975) or software (Pham 1985).

Typical forms of redundant hardware structures for fault-tolerant systems are of two types: fault masking and standby. Masking redundancy is achieved by implementing the functions so that they are inherently error correcting, *e.g.* triple-modular redundancy (TMR), N-modular redundancy (NMR), and self-purging redundancy. In standby redundancy, spare units are switched into the system when working units break down. Mathur and De Sousa (1975) have analyzed, in detail, hardware redundancy in the design of fault-tolerant digital systems. Redundant software structures for fault-tolerant systems based on the acceptance tests have been proposed by Homing *et al.* (1974).

This section presents a fault-tolerant architecture to increase the reliability of a special class of digital systems in communication (Pham and Upadhyaya 1989b). In this system, a monitor and a switch are associated with each redundant unit. The switches and monitors can fail. The monitors have two failure modes: failure to accept a correct result, and failure to reject an incorrect result. The scheme can be used in communication systems to improve their reliability.

Consider a digital circuit module designed to process the incoming messages in a communication system. This module consists of two units: a converter to process the messages, and a monitor to analyze the messages for their accuracy. For example, the converter could be decoding or unpacking circuitry, whereas the monitor could be checker circuitry (Lala 1985).

To guarantee a high reliability of operation at the receiver end, n converters are arranged in "parallel". All, except converter n, have a monitor to determine if the output of the converter is correct. If the output of a converter is not correct, the output is cancelled and a switch is changed so that the original input message is sent to the next converter. The architecture of such a system has been proposed by Pham and Upadhyaya (1989b). Systems of this kind have useful applications in communication and network control systems and in the analysis of fault-tolerant software systems.

We assume that a switch is never connected to the next converter without a signal from the monitor, and the probability that it is connected when a signal arrives is p_s. We next present a general expression for the reliability of the system consisting of n non-identical converters arranged in "parallel". Let us define the following notation, events, and assumptions.

Notation

p_i^c Pr{converter i works}

p_i^s Pr{switch i is connected to converter $(i + 1)$ when a signal arrives}

p_i^{m1} Pr{monitor i works when converter i works} = Pr{not sending a signal to the switch when converter i works}

p_i^{m2} Pr{i monitor works when converter i has failed} = Pr{sending a signal to the switch when converter i has failed}

R_{n-k}^k Reliability of the remaining system of size $(n-k)$ given that the first k switches work

R_n Reliability of the system consisting of n converters

The events are:

C_i^w, C_i^f Converter i works, fails

M_i^w, M_i^f Monitor i works, fails

S_i^w, S_i^f Switch i works, fails

W System works

The assumptions are:

1. The system, the switches, and the converters are two-state: good or failed.
2. The module (converter, monitor, or switch) states are mutually statistically independent.
3. The monitors have three states: good, failed in mode 1, failed in mode 2.
4. The modules are not identical.

The reliability of the system is defined as the probability of obtaining the correctly processed message at the output. To derive a general expression for the reliability of the system, we use an adapted form of the total probability theorem as translated into the language of reliability.

Let A denote the event that a system performs as desired. Let X_i and X_j be the event that a component X (*e.g.* converter, monitor, or switch) is good or failed respectively. Then

Pr{system works} = Pr{system works when unit X is good} x *Pr{unit X is good}*
+ Pr{system works when unit X fails} x *Pr{unit X is failed}*

The above equation provides a convenient way of calculating the reliability of complex systems. Notice that $R_1 = p_i^c$ and for $n \geq 2$, the reliability of the system can be calculated as follows:

$$R_n = \Pr\{W \mid C_1^w \text{ and } M_1^w\} \Pr\{C_1^w \text{ and } M_1^w\} + \Pr\{W \mid C_1^w \text{ and } M_1^f\} \Pr\{C_1^w \text{ and } M_1^f\} + \Pr\{W \mid C_1^f \text{ and } M_1^w\} \Pr\{C_1^f \text{ and } M_1^w\}$$

$$+\Pr\{W|C_1^f \text{ and } M_1^f\}\Pr\{C_1^f \text{ and } M_1^f\}$$

In order for the system to operate when the first converter works and the first monitor fails, the first switch must work and the remaining system of size n-1 must work:

$$\Pr\{W \mid C_1^w \text{ and } M_1^f\} = p_1^s R_{n-1}^1$$

Similarly,

$$\Pr\{W \mid C_1^f \text{ and } M_1^w\} = p_1^s R_{n-1}^1$$

then

$$R_n = p_1^c p_1^{m_1} + [p_1^c(1-p_1^{m_1}) + (1-p_1^c)p_1^{m_2}]p_1^s R_{n-1}^1$$

The reliability of the system consisting of n non-identical converters can be rewritten as:

$$R_n = \sum_{i=1}^{n-1} p_i^c p_i^{m1} \pi_{i-1} + \pi_{n-1} p_n^c \quad \text{for } n > 1 \tag{2.76}$$

and $R_1 = p_1^c$ where

$$\pi_k^j = \prod_{i=j}^{k} A_i \quad \text{for } k \geq 1$$

$$\pi_k \equiv \pi_k^1 \qquad \text{for all } k \text{ and } \pi_0 = 1$$

and

$$A_i \equiv [p_i^c(1-p_i^{m1}) + (1-p_i^c)p_i^{m2}] \quad \text{for all } i = 1, 2, ..., n$$

Assume that all the converters, monitors, and switches have the same reliability, that is

$$p_i^c = p^c, \quad p_i^{m1} = p^{m1}, \quad p_i^{m2} = p^{m2}, \quad p_i^s = p^s \quad \text{for all } i$$

then we obtain a closed form expression for the reliability of system as follows:

$$R_n = \frac{p^c p^{m1}}{1-A}(1-A^{n-1}) + p^c A^{n-1} \tag{2.77}$$

where

$$A = [p^c(1-p^{m1}) + (1-p^c)p^{m2}]p^s$$

2.5 Markov Processes

Stochastic processes are used for the description of a systems operation over time. There are two main types of stochastic processes: continuous and discrete. The complex continuous process is a process describing a system transition from state to state. The simplest process that will be discussed here is a Markov process. Given the current state of the process, its future behavior does not depend on the past. In Section 2.6 we will discuss the discrete stochastic process. As an introduction to the Markov process, let us examine the following example.

Example 2.15: Consider a parallel system consisting of two components. From a reliability point of view, the states of the system can be described by

State 1: Full operation (both components operating)

State 2: One component operating - one component failed

State 3: Both components failed

Define

$$P_i(t) = P[X(t) = i] = P[\text{system is in state } i \text{ at time } t]$$

and

$$P_i(t+dt) = P[X(t+dt) = i] = P[\text{system is in state } i \text{ at time } t+dt].$$

Define a random variable $X(t)$ which can assume the values 1, 2, or 3 corresponding to the above-mentioned states. Since $X(t)$ is a random variable, one can discuss $P[X(t) = 1]$, $P[X(t) = 2]$ or conditional probability, $P[X(t_1)= 2 \mid X(t_0)= 1]$. Again, $X(t)$ is defined as a function of time t, the last stated conditional probability, $P[X(t_1)= 2 \mid X(t_0)= 1]$, can be interpreted as the probability of being in state 2 at time t_1, given that the system was in state 1 at time t_0. In this example, the "stage space" is discrete, *i.e.*, 1, 2, 3, *etc.*, and the parameter space (time) is continuous. The simple process described above is called a stochastic process, *i.e.*, a process which develops in time (or space) in accordance with some probabilistic (stochastic) laws. There are many types of stochastic processes. In this section, the emphasis will be on Markov processes which are a special type of stochastic process.

Definition 2.3: Let $t_0 < t_1 < ... < t_n$. If

$$\begin{aligned} P[X(t_n) = A_n \mid X(t_{n-1}) = A_{n-1}, X(t_{n-2}) = A_{n-2}, ..., X(t_0) = A_0] \\ = P[X(t_n) = A_n \mid X(t_{n-1}) = A_{n-1}] \end{aligned}$$

then the process is called a Markov process.
Given the present state of the process, its future behavior does not depend on past information of the process.

The essential characteristic of a Markov process is that it is a process that has no memory; its future is determined by the present and not the past. If, in addition to having no memory, the process is such that it depends only on the difference $(t+dt)-t = dt$ and not the value of t, *i.e.*, $P[X(t + dt) = j \mid X(t) = i]$ is independent of t, then the process is Markov with stationary transition probabilities or homogeneous in time. This is the same property noted in exponential event times, and referring back to the graphical representation of $X(t)$, the times between state changes would in fact be exponential if the process has stationary transition probabilities.

Thus, a Markov process which is time homogeneous can be described as a process where events have exponential occurrence times. The random variable of the process is $X(t)$, the state variable rather than the time to failure as in the exponential failure density. To see the types of processes that can be described, a review of the exponential distribution and its properties will be made. Recall that, if X_1, X_2,…, X_n, are independent random variables, each with exponential density and a mean equal to $1/\lambda_i$ then min { $X_1, X_2, \ldots, X_n$} has an exponential density with mean $\left(\sum \lambda_i\right)^{-1}$.

The significance of the property is as follows:

1. The failure behavior of the simultaneous operation of components can be characterized by an exponential density with a mean equal to the reciprocal of the sum of the failure rates.
2. The joint failure/repair behavior of a system where components are operating and/or undergoing repair can be characterized by an exponential density with a mean equal to the reciprocal of the sum of the failure and repair rates.
3. The failure/repair behavior of a system such as 2 above, but further complicated by active and dormant operating states and sensing and switching, can be characterized by an exponential density.

The above property means that almost all reliability and availability models can be characterized by a time homogeneous Markov process if the various failure times and repair times are exponential. The notation for the Markov process is $\{X(t), t>0\}$, where $X(t)$ is discrete (state space) and *t is* continuous (parameter space). By convention, this type of Markov process is called a continuous parameter Markov chain.

From a reliability/availability viewpoint, there are two types of Markov processes. These are defined as follows:

1. *Absorbing Process:* Contains what is called an "absorbing state" which is a state from which the system can never leave once it has entered, *e.g.*, a failure which aborts a flight or a mission.
2. *Ergodic Process* Contains no absorbing states such that $X(t)$ can move around indefinitely, *e.g.*, the operation of a ground power plant where failure only temporarily disrupts the operation.

Pham (2000a) page 265, presents a summary of the processes to be considered broken down by absorbing and ergodic categories. Both reliability and availability can be described in terms of the probability of the process or system being in defined "up" states, e.g., states 1 and 2 in the initial example. Likewise, the mean time between failures (MTBF) can be described as the total time in the “up” states before proceeding to the absorbing state or failure state.

Define the incremental transition probability as

$$P_{ij}(dt) = P[X(t+dt) = j \mid X(t) = i]$$

This is the probability that the process (random variable $X(t)$) will go to state i during the increment t to $(t+dt)$, given that it was in state i at time t. Since we are dealing with time homogeneous Markov processes, *i.e.*, exponential failure and repair times, the incremental transition probabilities can be derived from an analysis of the exponential hazard function. In Section 2.1, it was shown that the hazard function for the exponential with mean $1/\lambda$ was just λ. This means that the limiting (as $dt \to 0$) conditional probability of an event occurrence between t and $t + dt$, given that an event had not occurred at time *t, is* just λ, *i.e.*,

$$h(t) = \lim_{dt \to 0} \frac{P[t < X < t + dt \mid X > t]}{dt} = \lambda$$

The equivalent statement for the random variable $X(t)$ is

$$h(t)dt = P[X(t + dt) = j \mid X(t) = i] = \lambda dt$$

Now, $h(t)dt$ is in fact the incremental transition probability, thus the $P_{ij}(dt)$ can be stated in terms of the basic failure and/or repair rates.

Returning to Example 2.15, a state transition can be easily constructed showing the incremental transition probabilities for process between all possible states (see Figure.B.4, Pham 2000a)

State 1: Both components operating

State 2: One component up - one component down

State 3: Both components down (absorbing state)

The loops indicate the probability of remaining in the present state during the *dt* increment

$P_{11}(dt) = 1 - 2\lambda dt$	$P_{12}(dt) = 2\lambda dt$	$P_{13}(dt) = 0$
$P_{21}(dt) = 0$	$P_{22}(dt) = 1 - \lambda dt$	$P_{23}(dt) = \lambda dt$
$P_{31}(dt) = 0$	$P_{32}(dt) = 0$	$P_{33}(dt) = 1$

The zeros on P_{ij}, $i > j$, denote that the process cannot go backwards, *i.e.*, this is not a repair process. The zero on P_{13} denotes that in a process of this type, the probability of more than one event (*e.g.*, failure, repair, *etc.*) in the incremental time period *dt* approaches zero as *dt* approaches zero.

Except for the initial conditions of the process, *i.e.*, the state in which the process starts, the process is completely specified by the incremental transition probabilities. The reason for the latter is that the assumption of exponential event (failure or repair) times allows the process to be characterized at any time *t* since it depends only on what happens between *t* and $(t+dt)$. The incremental transition probabilities can be arranged into a matrix in a way which depicts all possible statewide movements. Thus, for the parallel configurations,

$$[p_{ij}(dt)] = \begin{bmatrix} 1 & 2 & 3 \\ 1-2\lambda dt & 2\lambda dt & 0 \\ 0 & 1-\lambda dt & \lambda dt \\ 0 & 0 & 1 \end{bmatrix}$$

for $i, j = 1, 2$, or 3. The matrix $[P_{ij}(dt)]$ is called the incremental, one-step transition matrix. It is a stochastic matrix, *i.e.*, the rows sum to 1.0. As mentioned earlier, this matrix along with the initial conditions completely describes the process.

Now, $[P_{ij}(dt)]$ gives the probabilities for either remaining or moving to all the various states during the interval t to $t + dt$, hence,

$$P_1(t + dt) = (1 - 2\lambda dt)P_1(t)$$
$$P_2(t + dt) = 2\lambda dt\, P_1(t)(1 - \lambda dt)P_2(t)$$
$$P_3(t + dt) = \lambda dt P_2(t) + P_3(t)$$

By algebraic manipulation, we have

$$\frac{[P_1(t + dt) - P_1(t)]}{dt} = -2\lambda P_1(t)$$
$$\frac{[P_2(t + dt) - P_2(t)]}{dt} = 2\lambda P_1(t) - \lambda P_2(t)$$
$$\frac{[P_3(t + dt) - P_3(t)]}{dt} = \lambda P_2(t)$$

Taking limits of both sides as $dt \to 0$, we obtain

$$P_1'(t) = - 2\lambda P_1(t)$$
$$P_2'(t) = 2\lambda P_1(t) - 2\lambda P_2(t)$$
$$P_3'(t) = \lambda P_2(t)$$

The above system of linear first-order differential equations can be easily solved for $P_1(t)$ and $P_2(t)$, and therefore, the reliability of the configuration can be obtained:

$$R(t) = \sum_{i=1}^{2} P_i(t)$$

Actually, there is no need to solve all three equations, but only the first two as $P_3(t)$ does not appear and also $P_3(t) = 1 - P_1(t) - P_2(t)$. The system of linear, first-order differential equations can be solved by various means including both manual and machine methods. For purposes here, the manual methods employing the Laplace transform (see Appendix 2) will be used.

$$L[P_i(t)] = \int_0^\infty e^{-st} P_i(t)dt = f_i(s)$$

$$L[P_i^{'}(t)] = \int_0^\infty e^{-st} P_i^{'}(t)dt = s\, f_i(s) - P_i(0)$$

The use of the Laplace transform will allow transformation of the system of linear, first-order differential equations into a system of linear algebraic equations which can easily be solved, and by means of the inverse transforms, solutions of $P_i(t)$ can be determined.

Returning to the example, the initial condition of the parallel configuration is assumed to be "full-up" such that

$$P_1(t = 0) = 1,\ P_2(t = 0) = 0,\ P_3(t = 0) = 0$$

transforming the equations for $P'_1(t)$ and $P'_2(t)$ gives

$$sf_1(s) - P_1(t)\,|_{t=0} = -2\lambda f_1(s)$$
$$sf_2(s) - P_2(t)\,|_{t=0} = 2\lambda f_1(s) - \lambda f_2(s)$$

Evaluating $P_1(t)$ and $P_2(t)$ at $t = 0$ gives

$$sf_1(s) - 1 = -2\lambda f_1(s)$$
$$sf_2(s) - 0 = 2\lambda f_1(s) - \lambda f_2(s)$$

from which we obtain

$$(s + 2\lambda) f_1(s) = 1$$
$$-2\lambda f_1(s) + (s + \lambda) f_2(s) = 0$$

Solving the above equations for $f_1(s)$ and $f_2(s)$, we have

$$f_1(s) = \frac{1}{(s+2\lambda)}$$

$$f_2(s) = \frac{2\lambda}{[(s+2\lambda)(s+\lambda)]}$$

From Appendix 2 of the inverse Laplace transforms,

$$P_1(t) = e^{-2\lambda t}$$

$$P_2(t) = 2e^{-\lambda t} - 2e^{-2\lambda t}$$

$$R(t) = P_1(t) + P_2(t) = 2e^{-\lambda t} - e^{-2\lambda t}$$

The example given above is that of a simple absorbing process where we are concerned about reliability If repair capability in the form of a repair rate μ were added to the model, the methodology would remain the same with only the final result changing. With a repair rate μ added to the parallel configuration, the incremental transition matrix would be

$$[P_{ij}(dt)] = \begin{bmatrix} 1-2\lambda dt & 2\lambda dt & 0 \\ \mu dt & 1-(\lambda+\mu)dt & \lambda dt \\ 0 & 0 & 1 \end{bmatrix}$$

The differential equations would become

$$P_1'(t) = -2\lambda P_1(t) + \mu P_2(t)$$
$$P_2'(t) = 2\lambda P_1(t) - (\lambda + \mu)P_2(t)$$

and the transformed equations would become

$$(s+2\lambda)f_1(s) - \mu f_2(s) = 1$$

$$-2\lambda f_1(s) + (s+\lambda+\mu)f_2(s) = 0$$

Hence, we obtain

$$f_1(s) = \frac{(s+\lambda+\mu)}{(s-s_1)(s-s_2)}$$

$$f_2(s) = \frac{2\lambda}{(s-s_1)(s-s_2)}$$

where

$$s_1 = \frac{-(3\lambda+\mu)+\sqrt{(3\lambda+\mu)2-8\lambda^2}}{2}$$

$$s_2 = \frac{-(3\lambda+\mu)-\sqrt{(3\lambda+\mu)2-8\lambda^2}}{2}$$

From Appendix 2, we obtain

$$P_1(t) = \frac{(s_1 + \lambda + \mu)e^{-s_1 t}}{(s_1 - s_2)} + \frac{(s_2 + \lambda + \mu)e^{-s_2 t}}{(s_2 - s_1)}$$

$$P_2(t) = \frac{2\lambda e^{-s_1 t}}{(s_1 - s_2)} + \frac{2\lambda e^{-s_2 t}}{(s_2 - s_1)}$$

Thus, the reliability of two-component in a parallel system is given by

$$R(t) = P_1(t) + P_2(t)$$
$$= \frac{(s_1 + 3\lambda + \mu)e^{-s_1 t} - (s_2 + 3\lambda + \mu)e^{-s_2 t}}{(s_1 - s_2)} \tag{2.78}$$

System Mean Time Between Failures

Another parameter of interest in absorbing Markov processes is the mean time between failures (MTBF). Recalling the previous example of a parallel configuration with repair, the differential equations $P_1'(t)$ and $P_2'(t)$ describing the process were

$$P_1'(t) = -2\lambda P_1(t) + \mu P_2(t)$$
$$P_2'(t) = 2\lambda P_1(t) - (\lambda + \mu)P_2(t)$$

Integrating both sides of the above equations yields

$$\int_0^\infty P'_1(t)dt = -2\lambda \int_0^\infty P_1(t)dt + \mu \int_0^\infty P_2(t)dt$$

$$\int_0^\infty P_2'(t)dt = 2\lambda \int_0^\infty P_1(t)dt - (\lambda + \mu)\int_0^\infty P_2(t)dt$$

From Section 2.1,

$$\int_0^\infty R(t)dt = MTBF$$

Similarly,

$$\int_0^\infty P_1(t)dt = \text{ mean time spent in state 1, and}$$

$$\int_0^\infty P_2(t)dt = \text{ mean time spent in state 2}$$

Designating these mean times as T_1 and T_2, respectively, we have

$$P_1(t)dt \,|_0^\infty = -2\lambda T_1 + \mu T_2$$
$$P_2(t)dt \,|_0^\infty = 2\lambda T_1 - (\lambda + \mu)T_2$$

But $P_1(t) = 0$ as $t \to \infty$ and $P_1(t) = 1$ for $t = 0$. Likewise, $P_2(t) = 0$ as $t \to \infty$ and $P_2(t) = 0$ for $t = 0$. Thus,

$$-1 = -2\lambda T_1 + \mu T_2$$
$$0 = 2\lambda T_1 - (\lambda + \mu) T_2$$

or, equivalently,

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2\lambda & \mu \\ 2\lambda & -(\lambda+\mu) \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$$

Therefore,

$$T_1 = \frac{(\lambda+\mu)}{2\lambda^2} \qquad T_2 = \frac{1}{\lambda}$$

$$MTBF = T_1 + T_2 = \frac{(\lambda+\mu)}{2\lambda^2} + \frac{1}{\lambda} = \frac{(3\lambda+\mu)}{2\lambda^2}$$

The MTBF for non-maintenance processes is developed exactly the same way as just shown. What remains under absorbing processes is the case for availability for maintained systems. The difference between reliability and availability for absorbing processes is somewhat subtle. A good example is that of a communication system where, if such a system failed temporarily, the mission would continue, but, if it failed permanently, the mission would be aborted. Consider the following cold-standby configuration consisting of two units: one main unit and one spare unit (Pham 2000a):

State 1: Main unit operating - spare OK
State 2: Main unit out - restoration underway
State 3: Spare unit installed and operating
State 4: Permanent failure (no spare available)

The incremental transition matrix is given by (see Figure B.8 in Pham 2000a, for a detailed state transition diagram)

$$[P_{ij}(dt)] = \begin{bmatrix} 1-\lambda dt & \lambda dt & 0 & 0 \\ 0 & 1-\mu dt & \mu dt & 0 \\ 0 & 0 & 1-\lambda dt & \lambda dt \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We obtain

$$P_1'(t) = -\lambda P_1(t)$$
$$P_2'(t) = \lambda P_1(t) - \mu P_2(t)$$
$$P_3'(t) = \mu P_2(t) - \lambda P_3(t)$$

Using the Laplace transform, we obtain

$$sf_1(s) - 1 = -\lambda f_1(s)$$
$$sf_2(s) = \lambda f_1(s) - \mu f_2(s)$$
$$sf_3(s) = \mu f_2(s) - \lambda f_3(s)$$

After simplifications,

$$f_1(s) = \frac{1}{(s+\lambda)}$$

$$f_2(s) = \frac{\lambda}{[(s+\lambda)(s+\mu)]}$$

$$f_3(s) = \frac{\lambda\mu}{[(s+\lambda)^2(s+\mu)]}$$

Therefore, the probability of full-up performance, $P_1(t)$, is given by

$$P_1(t) = e^{-\lambda t}$$

Similarly, the probability of the system being down and under repair, $P_2(t)$, is

$$P_2(t) = \left[\frac{\lambda}{(\lambda-\mu)}\right]\left(e^{-\mu t} - e^{-\lambda t}\right)$$

and the probability of the system being full-up but no spare available, $P_3(t)$, is

$$P_3(t) = \left[\frac{\lambda\mu}{(\lambda-\mu)^2}\right][e^{-\mu t} - e^{-\lambda t} - (\lambda-\mu)te^{-\lambda t}]$$

Hence, the point availability, $A(t)$, is given by

$$A(t) = P_1(t) + P_3(t)$$

If average or interval availability is required, this is achieved by

$$\left(\frac{1}{t}\right)\int_0^T A(t)dt = \left(\frac{1}{t}\right)\int_0^T [P_1(t) + P_3(t)]dt \tag{2.79}$$

where T is the interval of concern.

With the above example, cases of the absorbing process (both maintained and non-maintained) have been covered insofar as "manual" methods are concerned. In general, the methodology for treatment of absorbing Markov processes can be “packaged” in a fairly simplified form by utilizing matrix notation. Thus, for example, if the incremental transition matrix is defined as follows:

$$[P_{ij}(dt)] = \begin{bmatrix} 1\text{-}2\lambda dt & 2\lambda dt & 0 \\ \mu dt & 1-(\lambda+\mu)dt & \lambda dt \\ 0 & 0 & 1 \end{bmatrix}$$

then if the dts are dropped and the last row and the last column are deleted, the remainder is designated as the matrix T:

$$[T] = \begin{bmatrix} 1\text{-}2\lambda & 2\lambda \\ \mu & 1-(\lambda+\mu) \end{bmatrix}$$

Define $[Q] = [T]' - [I]$, where $[T]'$ is the transposition of $[T]$ and $[I]$ is the unity matrix:

$$[Q] = \begin{bmatrix} 1-2\lambda & \mu \\ 2\lambda & 1-(\lambda+\mu) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -2\lambda & \mu \\ 2\lambda & -(\lambda+\mu) \end{bmatrix}$$

Further define $[P(t)]$ and $[P'(t)]$ as column vectors such that

$$[P_1(t)] = \begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix}, \quad [P'(t)] = \begin{bmatrix} P_1'(t) \\ P_2'(t) \end{bmatrix}$$

then

$$[P'(t)] = [Q][P(t)]$$

At the above point, solution of the system of differential equations will produce solutions to $P_1(t)$ and $P_2(t)$. If the MTBF is desired, integration of both sides of the system produces

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} = [Q]\begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$$
$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2\lambda & \mu \\ 2\lambda & -(\lambda+\mu) \end{bmatrix}\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \quad \text{or}$$
$$[Q]^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$$

where $[Q]^{-1}$ is the inverse of $[Q]$ and the MTBF is given by

$$\text{MTBF} = T_1 + T_2 = \frac{3\lambda+\mu}{(2\lambda)^2}$$

In the more general MTBF case,

$$[Q]^{-1}\begin{bmatrix} -1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ \cdot \\ \cdot \\ \cdot \\ T_{n-1} \end{bmatrix} \quad \text{where } \sum_{i=1}^{n-1} T_i = \text{MTBF}$$

and $(n - 1)$ is the number of non-absorbing states.

For the reliability/availability case, utilizing the Laplace transform, the system of linear, first-order differential equations is transformed to

$$s\begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = [Q]\begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix}$$

$$[sI - Q]\begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix} = [sI - Q]^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$L^{-1}\begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix} = L^{-1}\{[sI - Q]^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}$$

$$\begin{bmatrix} p_1(s) \\ p_2(s) \end{bmatrix} = L^{-1}\{[sI - Q]^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}$$

Generalization of the latter to the case of (n-1) non-absorbing states is straightforward.

Ergodic processes, as opposed to absorbing processes, do not have any absorbing states, and hence, movement between states can go on indefinitely For the latter reason, availability (point, steady-state, or interval) is the only meaningful measure. As an example for ergodic processes, a ground-based power unit configured in parallel will be selected.

The parallel units are identical, each with exponential failure and repair times with means $1/\lambda$ and $1/\mu$, respectively (Pham 2000a). Assume a two-repairmen capability if required (both units down), then

State 1: Full-up (both units operating)
State 2: One unit down and under repair (other unit up)
State 3: Both units down and under repair

It should be noted that, as in the case of failure events, two or more repairs cannot be made in the *dt* interval.

$$[P_{ij}(dt)] = \begin{bmatrix} 1-2\lambda dt & 2\lambda dt & 0 \\ \mu dt & 1-(\lambda+\mu)dt & \lambda dt \\ 0 & 2\mu dt & 1-2\mu dt \end{bmatrix}$$

Case I: Point Availability - Ergodic Process. For an ergodic process, as $t \to \infty$ the availability settles down to a constant level. Point availability gives a measure of things before the "settling down" and reflects the initial conditions on the process. Solution of the point availability is similar to the case for absorbing processes except that the last row and column of the transition matrix must be retained and entered into the system of equations. For example, the system of differential equations becomes

$$\begin{bmatrix} P_1'(t) \\ P_2'(t) \\ P_3'(t) \end{bmatrix} = \begin{bmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \\ 0 & \lambda & -2\mu \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{bmatrix}$$

Similar to the absorbing case, the method of the Laplace transform can be used to solve for $P_1(t)$, $P_2(t)$, and $P_3(t)$, with the point availability, $A(t)$, given by

$$A(t) = P_1(t) + P_2(t)$$

Case II: Interval Availability - Ergodic Process. This is the same as the absorbing case with integration over time period T of interest. The interval availability, $A(T)$, is

$$A(T) = \frac{1}{T}\int_0^T A(t)dt \tag{2.80}$$

Case III: Steady State Availability - Ergodic Process. Here the process is examined as $t \to \infty$ with complete "washout" of the initial conditions. Letting $t \to \infty$ the system of differential equations can be transformed to linear algebraic equations. Thus,

$$\lim_{t\to\infty} \begin{bmatrix} P_1'(t) \\ P_2'(t) \\ P_3'(t) \end{bmatrix} = \lim_{t\to\infty} \begin{bmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \\ 0 & \lambda & -2\mu \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{bmatrix}$$

As $t \to \infty$, $P_i(t) \to$ constant and $P_i'(t) \to 0$. This leads to an unsolvable sys-tem, namely

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \\ 0 & \lambda & -2\mu \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{bmatrix}$$

To avoid the above difficulty, an additional equation is introduced:

$$\sum_{i=1}^{3} P_i(t) = 1$$

With the introduction of the new equation, one of the original equations is deleted and a new system is formed:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \end{bmatrix} \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{bmatrix}$$

or, equivalently,

$$\begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -2\lambda & \mu & 0 \\ 2\lambda & -(\lambda+\mu) & 2\mu \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

We now obtain the following results:

$$P_1(t) = \frac{\mu^2}{(\mu+\lambda)^2}$$

$$P_2(t) = \frac{2\lambda\mu}{(\mu+\lambda)^2}$$

and

$$P_3(t) = 1 - P_1(t) - P_2(t)$$
$$= \frac{\lambda^2}{(\mu+\lambda)^2}$$

Therefore, the steady state availability, $A(\infty)$, is given by

$$A_3(\infty) = P_1(t) + P_2(t)$$
$$= \frac{\mu(\mu+2\lambda)}{(\mu+\lambda)^2}$$

Note that Markov methods can also be employed where failure or repair times are not exponential, but can be represented as the sum of exponential times with identical means (Erlang distribution or Gamma distribution with integer valued shape parameters). Basically, the method involves the introduction of "dummy" states which are of no particular interest in themselves, but serve the purpose of changing the hazard function from constant to increasing.

2.6 Counting Processes

Among discrete stochastic processes, counting processes in reliability engineering are widely used to describe the appearance of events in time, *e.g.*, failures, number of perfect repairs, *etc*. The simplest counting process is a Poisson process. The Poisson process plays a special role to many applications in reliability (Pham 2000a). A classic example of such an application is the decay of uranium. Radioactive particles from nuclear material strike a certain target in accordance with a Poisson process of some fixed intensity. A well-known counting process is the so-called renewal process. This process is described as a sequence of events, the intervals between which are independent and identically distributed random variables. In reliability theory, this type of mathematical model is used to describe the number of occurrences of an event in the time interval. In this section we also discuss the quasi-renewal process and the non-homogeneous Poisson process.

A non-negative, integer-valued stochastic process, $N(t)$, is called a counting process if $N(t)$ represents the total number of occurrences of the event in the time interval [0, t] and satisfies these two properties:

1. If $t_1 < t_2$, then $N(t_1) \leq N(t_2)$
2. If $t_1 < t_2$, then $N(t_2) - N(t_1)$ is the number of occurrences of the event in the interval $[t_1, t_2]$

For example, if $N(t)$ equals the number of persons who have entered a restaurant at or prior to time t, then $N(t)$ is a counting process in which an event occurs whenever a person enters the restaurant.

2.6.1 Poisson Processes

One of the most important counting processes is the Poisson process.

Definition 2.4: A counting process, $N(t)$, is said to be a Poisson process with intensity λ if

1. The failure process, $N(t)$, has stationary independent increments
2. The number of failures in any time interval of length s has a Poisson distribution with mean λs, that is,

$$P\{N(t+s)-N(t)=n\}=\frac{e^{-\lambda s}(\lambda s)^n}{n!} \qquad n=0,1,2,... \tag{2.81}$$

3. The initial condition is $N(0) = 0$

This model is also called a homogeneous Poisson process indicating that the failure rate λ does not depend on time t. In other words, the number of failures occurring during the time interval $(t, t + s]$ does not depend on the current time t but only the length of time interval s. A counting process is said to possess independent increments if the number of events in disjoint time intervals are independent.

For a stochastic process with independent increments, the auto-covariance function is

$$Cov[X(t_1), X(t_2)] = \begin{cases} Var[N(t_1+s)-N(t_2)] & \text{for } 0<t_2-t_1<s \\ 0 & \text{otherwise} \end{cases}$$

where

$$X(t) = N(t+s) - N(t).$$

If $X(t)$ is Poisson distributed, then the variance of the Poisson distribution is

$$Cov[X(t_1), X(t_2)] = \begin{cases} \lambda[s-(t_2-t_1)] & \text{for } 0<t_2-t_1<s \\ 0 & \text{otherwise} \end{cases}$$

This result shows that the Poisson increment process is covariance stationary. We now present several properties of the Poisson process.

Property 2.3: The sum of independent Poisson processes, $N_1(t)$, $N_2(t)$, …., $N_k(t)$, with mean values $\lambda_1 t$, $\lambda_2 t$, …., $\lambda_k t$ respectively, is also a Poisson process with mean $\left(\sum_{i=1}^{k}\lambda_i\right)t$. In other words, the sum of the independent Poisson processes is also a Poisson process with a mean that is equal to the sum of the individual Poisson process' mean.

Proof: The proof is left as an exercise for the reader (see Problem 26).

Property 2.4: The difference of two independent Poisson processes, $N_1(t)$, and $N_2(t)$, with mean $\lambda_1 t$ and $\lambda_2 t$, respectively, is not a Poisson process. Instead, it has the probability mass function

$$P[N_1(t) - N_2(t) = k] = e^{-(\lambda_1+\lambda_2)t} \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{k}{2}} I_k(2\sqrt{\lambda_1\lambda_2}t) \tag{2.82}$$

where $I_k(.)$ is a modified Bessel function of order k (Handbook 1980).

Proof: The proof is left as an exercise for the reader (see Problem 27).

Property 2.5: If the Poisson process, $N(t)$, with mean λt, is filtered such that every occurrence of the event is not completely counted, then the process has a constant probability p of being counted. The result of this process is a Poisson process with mean λpt [ ].

Property 2.6: Let $N(t)$ be a Poisson process and Y_n a family of independent and identically distributed random variables which are also independent of $N(t)$. A stochastic process $X(t)$ *is* said to be a compound Poisson process if it can be represented as

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

2.6.2 Renewal Processes

A renewal process is a more general case of the Poisson process in which the inter-arrival times of the process or the time between failures do not necessarily follow the exponential distribution. For convenience, we will call the occurrence of an event a renewal, the inter-arrival time the renewal period, and the waiting time the renewal time.

Definition 2.5: A counting process $N(t)$ that represents the total number of occurrences of an event in the time interval $(0, t]$ is called a renewal process, if the time between failures are independent and identically distributed random variables.

The probability that there are exactly n failures occurring by time t can be written as

$$P\{N(t) = n\} = P\{N(t) \geq n\} - P\{N(t) > n\} \tag{2.83}$$

Note that the times between the failures are T_1, T_2, …, T_n so the failures occurring at time W_k are

$$W_k = \sum_{i=1}^{k} T_i$$

and

$$T_k = W_k - W_{k-1}$$

Thus,

$$\begin{aligned} P\{N(t)=n\} &= P\{N(t)\geq n\}-P\{N(t)>n\} \\ &= P\{W_n \leq t\}-P\{W_{n+1}\leq t\} \\ &= F_n(t)-F_{n+1}(t) \end{aligned}$$

where $F_n(t)$ *is* the cumulative distribution function for the time of the *n*th failure and $n = 0,1,2, \ldots$.

Example 2.16: Consider a software testing model for which the time to find an error during the testing phase has an exponential distribution with a failure rate of *X*. It can be shown that the time of the *n*th failure follows the gamma distribution with parameters *k* and *n* with probability density function. From equation (2.83) we obtain

$$\begin{aligned} P\{N(t)=n\} &= P\{N(t)\leq n\}-P\{N(t)\leq n-1\} \\ &= \sum_{k=0}^{n}\frac{(\lambda t)^k}{k!}e^{-\lambda t}-\sum_{k=0}^{n-1}\frac{(\lambda t)^k}{k!}e^{-\lambda t} \\ &= \frac{(\lambda t)^n}{n!}e^{-\lambda t} \quad \text{for} \quad n=0,1,2,.... \end{aligned}$$

Several important properties of the renewal function are given below.

Property 2.7: The mean value function of the renewal process, denoted by $m(t)$, is equal to the sum of the distribution function of all renewal times, that is,

$$\begin{aligned} m(t) &= E[N(t)] \\ &= \sum_{n=1}^{\infty} F_n(t) \end{aligned}$$

Proof: The renewal function can be obtained as

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

The mean value function of the renewal process is also called the renewal function.

Property 2.8: The renewal function, $m(t)$, satisfies the following equation:

$$m(t) = F_a(t) + \int_0^t m(t-s)dF_a(s) \tag{2.84}$$

where $F_a(t)$ is the distribution function of the inter-arrival time or the renewal period. The proof is left as an exercise for the reader (see Problem 28).

In general, let $y(t)$ be an unknown function to be evaluated and $x(t)$ be any non-negative and integrable function associated with the renewal process. Assume that $F_a(t)$ is the distribution function of the renewal period. We can then obtain the following result.

Property 2.9: Let the renewal equation be

$$y(t) = x(t) + \int_0^t y(t-s)dF_a(s) \tag{2.85}$$

then its solution is given by

$$y(t) = x(t) + \int_0^t x(t-s)dm(s)$$

where $m(t)$ is the mean value function of the renewal process.

The proof of the above property can be easily derived using the Laplace transform. It is also noted that the integral equation given in Property 2.8 is a special case of Property 2.9.

Example 2.17: Let $x(t) = a$. Thus, in Property 2.9, the solution $y(t)$ is given by

$$\begin{aligned} y(t) &= x(t) + \int_0^t x(t-s)dm(s) \\ &= a + \int_0^t a\ dm(s) \\ &= a(1 + E[N(t)]) \end{aligned}$$

2.6.3 Quasi-renewal Processes

In this section, a general renewal process, namely, the quasi-renewal process, is discussed. Let $\{N(t), t > 0\}$ be a counting process and let X_n be the time between the $(n\text{-}1)^{th}$ and the n^{th} event of this process, $n \geq 1$.

Definition 2.6 (Wang and Pham 1996a): If the sequence of non-negative random variables $\{X_1, X_2, ...\}$ is independent and

$$X_i = aX_{i-1} \tag{2.86}$$

for $i \geq 2$ where $\alpha > 0$ is a constant, then the counting process $\{N(t), t \geq 0\}$ is said to be a quasi-renewal process with parameter and the first inter-arrival time X_1.

When $\alpha = 1$, this process becomes the ordinary renewal process as discussed in Section 2.6.2. This quasi-renewal process can be used to model reliability growth processes in software testing phases and hardware burn-in stages for $\alpha > 1$, and in hardware maintenance processes when $\alpha \leq 1$.

Assume that the probability density function, cumulative distribution function, survival function, and failure rate of random variable X_1 are $f_1(x)$, $F_1(x)$, $s_1(x)$, and $r_1(x)$, respectively. Then the pfd, cdf, survival function, failure rate of X_n for $n = 1$, 2, 3,... is respectively given below (Wang and Pham 1996a):

$$f_n(x) = \frac{1}{\alpha^{n-1}} f_1\left(\frac{1}{\alpha^{n-1}} x\right)$$

$$F_n(x) = F_1\left(\frac{1}{\alpha^{n-1}} x\right)$$

$$s_n(x) = s_1\left(\frac{1}{\alpha^{n-1}} x\right)$$

$$f_n(x) = \frac{1}{\alpha^{n-1}} r_1\left(\frac{1}{\alpha^{n-1}} x\right)$$

Similarly, the mean and variance of X_n is given as

$$E(X_n) = \alpha^{n-1} E(X_1)$$

$$Var(X_n) = \alpha^{2n-2} Var(X_1)$$

Because of the non-negativity of X_1 and the fact that X_1 is not identically 0, we obtain

$$E(X_1) = \mu_1 \neq 0$$

Proposition 2.1 (Wang and Pham 1996a): The shape parameters of X_n are the same for $n = 1, 2, 3, \ldots$ for a quasi-renewal process if X_1 follows the gamma, Weibull, or log normal distribution.

This means that after "renewal", the shape parameters of the inter-arrival time will not change. In software reliability, the assumption that the software debugging process does not change the error-free distribution type seems reasonable. Thus, the error-free times of software during the debugging phase modeled by a quasi-renewal process will have the same shape parameters. In this sense, a quasi-renewal process is suitable to model the software reliability growth. It is worthwhile to note that

$$\begin{aligned}
\lim_{n\to\infty} \frac{E(X_1 + X_2 + \ldots + X_n)}{n} &= \lim_{n\to\infty} \frac{\mu_1(1-\alpha^n)}{(1-\alpha)n} \\
&= 0 \quad \text{if } \alpha < 1 \\
&= \infty \quad \text{if } \alpha > 1
\end{aligned}$$

Therefore, if the inter-arrival time represents the error-free time of a software system, then the average error-free time approaches infinity when its debugging process is occurring for a long debugging time.

Distribution of N(t)

Consider a quasi-renewal process with parameter α and the first inter-arrival time X_1. Clearly, the total number of renewals, $N(t)$, that has occurred up to time t and the arrival time of the nth renewal, SS_n, has the following relationship:

$$N(t) \geq n \text{ if and only if } SS_n \leq t$$

that is, $N(t)$ is at least n if and only if the nth renewal occurs prior to time t. It is easily seen that

$$SS_n = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \alpha^{i-1} X_1 \quad \text{for} \quad n \geq 1 \tag{2.87}$$

Here, $SS_0 = 0$. Thus, we have

$$\begin{aligned} P\{N(t)=n\} &= P\{N(t)=n\} - P\{N(t) \geq n+1\} \\ &= P\{SS_n \leq t\} - P\{SS_{n+1} \leq t\} \\ &= G_n(t) - G_{n+1}(t) \end{aligned}$$

where $G_n(t)$ is the convolution of the inter-arrival times F_1, F_2, F_3, …, F_n. In other words,

$$G_n(t) = P\{F_1 + F_2 + ... + F_n \leq t\}$$

If the mean value of $N(t)$ is defined as the renewal function $m(t)$, then,

$$\begin{aligned} m(t) &= E[N(t)] \\ &= \sum_{n=1}^{\infty} P\{N(t) \geq n\} \\ &= \sum_{n=1}^{\infty} P\{SS_n \leq t\} \\ &= \sum_{n=1}^{\infty} G_n(t) \end{aligned}$$

The derivative of $m(t)$ is known as the renewal density

$$\lambda(t) = m'(t)$$

In renewal theory, random variables representing the inter-arrival distributions only assume non-negative values, and the Laplace transform of its distribution $F_1(t)$ is defined by

$$\mathfrak{L}\{F_1(s)\} = \int_0^{\infty} e^{-sx} dF_1(x)$$

Therefore,

$$\mathfrak{L}F_n(s) = \int_0^{\infty} e^{-a^{n-1}st} dF_1(t) = \mathfrak{L}F_1(\alpha^{n-1}s)$$

and

$$\mathfrak{L}m_n(s) = \sum_{n=1}^{\infty} \mathfrak{L}G_n(s)$$

$$= \sum_{n=1}^{\infty} \mathfrak{L}F_1(s)\mathfrak{L}F_1(\alpha s) \cdots \mathfrak{L}F_1(\alpha^{n-1}s)$$

Since there is a one-to-one correspondence between distribution functions and its Laplace transform, it follows that

Proposition 2.2 (Wang and Pham 1996a): The first inter-arrival distribution of a quasi-renewal process uniquely determines its renewal function.

If the inter-arrival time represents the error-free time (time to first failure), a quasi-renewal process can be used to model reliability growth for both software and hardware.

Suppose that all faults of software have the same chance of being detected. If the inter-arrival time of a quasi-renewal process represents the error-free time of a software system, then the expected number of software faults in the time interval [0, t] can be defined by the renewal function, $m(t)$, with parameter $\alpha > 1$. Denoted by $m_r(t)$, the number of remaining software faults at time t, it follows that

$$m_r(t) = m(T_c) - m(t)$$

where $m(T_c)$ *is* the number of faults that will eventually be detected through a software lifecycle T_c.

2.6.4 Non-homogeneous Poisson Processes

The non-homogeneous Poisson process model (NHPP) that represents the number of failures experienced up to time t is a non-homogeneous Poisson process $\{N(t), t \geq 0\}$. The main issue in the NHPP model is to determine an appropriate mean value function to denote the expected number of failures experienced up to a certain time.

With different assumptions, the model will end up with different functional forms of the mean value function. Note that in a renewal process, the exponential assumption for the inter-arrival time between failures is relaxed, and in the NHPP, the stationary assumption is relaxed.

The NHPP model is based on the following assumptions:

- The failure process has an independent increment, *i.e.*, the number of failures during the time interval (t, $t + s$) depends on the current time t and the length of time interval s, and does not depend on the past history of the process.
- The failure rate of the process is given by

  $$P\{\text{exactly one failure in } (t, t+\Delta t)\} = P\{N(t+\Delta t) - N(t) = 1\}$$
  $$= \lambda(t)\Delta t + o(\Delta t)$$

  where $\lambda(t)$ is the intensity function.
- During a small interval Δt, the probability of more than one failure is negligible, that is,

$P\{\text{two or more failure in } (t, t+\Delta t)\} = o(\Delta t)$

- The initial condition is $N(0) = 0$.

On the basis of these assumptions, the probability of exactly n failures occurring during the time interval (0, t) for the NHPP is given by

$$\Pr\{N(t) = n\} = \frac{[m(t)]^n}{n!} e^{-m(t)} \qquad n = 0, 1, 2, \ldots \tag{2.88}$$

where $m(t) = E[N(t)] = \int_0^t \lambda(s)ds$ and $\lambda(t)$ is the intensity function. It can be easily shown that the mean value function $m(t)$ is non-decreasing.

Reliability Function

The reliability $R(t)$, defined as the probability that there are no failures in the time interval (0, t), is given by

$$\begin{aligned} R(t) &= P\{N(t) = 0\} \\ &= e^{-m(t)} \end{aligned}$$

In general, the reliability $R(x|t)$, the probability that there are no failures in the interval (t, $t + x$), is given by

$$\begin{aligned} R(x \mid t) &= P\{N(t+x) - N(t) = 0\} \\ &= e^{-[m(t+x) - m(t)]} \end{aligned}$$

and its density is given by

$$f(x) = \lambda(t+x) e^{-[m(t+x) - m(t)]}$$

where

$$\lambda(x) = \frac{\partial}{\partial x}[m(x)]$$

The variance of the NHPP can be obtained as follows:

$$Var[N(t)] = \int_0^t \lambda(s)ds$$

and the auto-correlation function is given by

$$\begin{aligned} Cor[s] &= E[N(t)]E[N(t+s) - N(t)] + E[N^2(t)] \\ &= \int_0^t \lambda(s)ds \int_0^{t+s} \lambda(s)ds + \int_0^t \lambda(s)ds \\ &= \int_0^t \lambda(s)ds \left[1 + \int_0^{t+s} \lambda(s)ds \right] \end{aligned}$$

Example 2.18: Assume that the intensity λ is a random variable with the pdf $f(\lambda)$. Then the probability of exactly n failures occurring during the time interval (0, t) is given by

$$P\{N(t) = n\} = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} f(\lambda) d\lambda$$

It can be shown that if the pdf $f(\lambda)$ is given as the following gamma density function with parameters k and m,

$$f(\lambda) = \frac{1}{\Gamma(m)} k^m \lambda^{m-1} e^{-k\lambda} \quad \text{for } \lambda \geq 0$$

then

$$P\{N(t) = n\} = \binom{n+m-1}{n} p^m q^n \quad n = 0, 1, 2, ...$$

is also called a negative binomial density function, where

$$p = \frac{k}{t+k} \quad \text{and} \quad q = \frac{t}{t+k} = 1 - p$$

2.7 Further Reading

The reader interested in a deeper understanding of advanced probability theory and stochastic processes should note the following highly recommended books:

Devore, J.L., *Probability and Statistics for Engineering and the Sciences,* 3rd edition, Brooks/Cole Pub. Co., Pacific Grove, 1991.
Gnedenko, BV and I.A. Ushakov, *Probabilistic Reliability Engineering*, Wiley, New York, 1995.
Feller, W., *An Introduction to Probability Theory and Its Applications*, 3rd edition, Wiley, New York, 1994.

2.8 Problems

1. Assume that the hazard rate, $h(t)$, has a positive derivative. Show that the hazard distribution

$$H(t) = \int_0^t h(x) dx$$

is strictly convex.

2. An operating unit is supported by (n-1) identical units on cold standby. When it fails, a unit from standby takes its place. The system fails if all n units fail.

Assume that units on standby cannot fail and the lifetime of each unit follows the exponential distribution with failure rate λ.

(a) What is the distribution of the system lifetime?

(b) Determine the reliability of the standby system for a mission of 100 hours when λ = 0.0001 per hour and n = 5.

3. Assume that there is some latent deterioration process occurring in the system. During the interval [0, *a-h*] the deterioration is comparatively small so that the shocks do not cause system failure. During a relatively short time interval [*a-h, a*], the deterioration progresses rapidly and makes the system susceptible to shocks. Assume that the appearance of each shock follows the exponential distribution with failure rate λ. What is the distribution of the system lifetime?

4. Consider a series system of n Weibull components. The corresponding lifetimes $T_1, T_2, \ldots, T_n$ are assumed to be independent with pdf

$$f(t) = \begin{cases} \lambda_i^{\beta} \beta t^{\beta-1} e^{-(\lambda_i t)^{\beta}} & \text{for } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda > 0$ and $\beta > 0$ are the scale and shape parameters, respectively.

(a) Show that the lifetime of a series system has the Weibull distribution with pdf

$$f_s(t) = \begin{cases} \left(\sum_{i=1}^{n} \lambda_i^{\beta}\right) \beta t^{\beta-1} e^{-\left(\sum_{i=1}^{n} \lambda_i^{\beta}\right) t^{\beta}} & \text{for } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) Find the reliability of this series system.

5. Consider the pdf of a random variable that is equally likely to take on any value *only* in the interval from a to b.

(a) Show that this pdf is given by

$$f(t) = \begin{cases} \dfrac{1}{b-a} & \text{for } a < t < 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) Derive the corresponding reliability function $R(t)$ and failure rate $h(t)$.

(c) Think of an example where such a distribution function would be of interest in reliability application.

6. The failure rate function, denoted by $h(t)$, is defined as

$$h(t) = -\frac{d}{dt} \ln[R(t)]$$

Show that the constant failure rate function implies an exponential distribution.

7. One thousand new streetlights are installed in Saigon city. Assume that the lifetimes of these streetlights follow the normal distribution. The average life of these lamps is estimated at 980 burning-hours with a standard deviation of 100 hours.

(a) What is the expected number of lights that will fail during the first 800 burning-hours?

(b) What is the expected number of lights that will fail between 900 and 1100 burning-hours?

(c) After how many burning-hours would 10% of the lamps be expected to fail?

8. A fax machine with constant failure rate λ will survive for a period of 720 hours without failure, with probability 0.80.

(a) Determine the failure rate λ.

(b) Determine the probability that the machine, which is functioning after 600 hours, will still function after 800 hours.

(c) Find the probability that the machine will fail within 900 hours, given that the machine was functioning at 720 hours.

9. The time to failure T of a unit is assumed to have a log normal distribution with pdf

$$f(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma t} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} \qquad t > 0$$

Show that the failure rate function is unimodal.

10. A diode may fail due to either open or short failure modes. Assume that the time to failure T_0 caused by open mode is exponentially distributed with pdf

$$f_0(t) = \lambda_0 e^{-\lambda_0 t} \qquad t \geq 0$$

and the time to failure T_1 caused by short mode has the pdf

$$f_s(t) = \lambda_s e^{-\lambda_s t} \qquad t \geq 0$$

The pdf for the time to failure T of the diode is given by

$$f(t) = p\, f_0(t) + (1-p) f_s(t) \qquad t \geq 0$$

(a) Explain the meaning of p in the above pdf function.

(b) Derive the reliability function $R(t)$ and failure rate function $h(t)$ for the time to failure T of the diode.

(c) Show that the diode with pdf $f(t)$ has a decreasing failure rate (DFR).

11. A diesel is known to have an operating life (in hours) that fits the following pdf:

$$f(t) = \frac{2a}{(t+b)^2} \qquad t \geq 0$$

The average operating life of the diesel has been estimated to be 8760 hours.

(a) Determine a and b.

(b) Determine the probability that the diesel will not fail during the first 6000 operating-hours.

(c) If the manufacturer wants no more than 10% of the diesels returned for warranty service, how long should the warranty be?

12. The failure rate for a hydraulic component

$$h(t) = \frac{t}{t+1} \qquad t > 0$$

where t is in years.

(a) Determine the reliability function $R(t)$.

(b) Determine the MTTF of the component.

13. A 18-month guarantee is given based on the assumption that no more than 5% of new cars will be returned.

(a) The time to failure T of a car has a constant failure rate. What is the maximum failure rate that can be tolerated?

(b) Determine the probability that a new car will fail within three years assuming that the car was functioning at 18 months.

14. Show that if

$$R_1(t) \geq R_2(t) \quad \text{for all } t$$

where $R_i(t)$ *is* the system reliability of the structure i, then MTTF of the system structure 1 is always ≥ MTTF of the system structure 2.

15. Prove equation (2.10)

16. Show that the reliability function of Pham distribution (see equation 2.21) is given as in equation (2.22).

17. Prove Theorem 2.1.

18. Prove Theorem 2.2.

19. Show that for any range of q_0 and q_s, if $q_s > q_0$, the optimal number of parallel components that maximizes the system reliability is one.

20. Prove Theorem 2.3.

21. Prove Theorem 2.4.

22. Prove Theorem 2.5.

23. Show that the optimal value n^* in Theorem 2.5 is an increasing function of q_0 and a decreasing function of q_s.

24. Prove Theorem 2.6.

25. For any given q_0 and q_s, show that $q_s < \alpha < p_0$ where α is given in equation (2.73).

26. Prove Property 2.3.

27. Prove Property 2.4.

28. Prove Property 2.8.

29. Events occur according to an NHPP in which the mean value function is

$$m(t) = t^3 + 3t^2 + 6t \quad t > 0.$$

What is the probability that n events occur between times $t = 10$ and $t = 15$?

Springer

http://www.springer.com/978-1-85233-950-0

System Software Reliability
Pham, H.
2006, XIV, 440 p., Hardcover
ISBN: 978-1-85233-950-0