Maintenance modeling and optimization

Research School for Operations Management and Logistics

Maintenance Modeling and Optimization

J.J. Arts

Beta Working Paper series 526

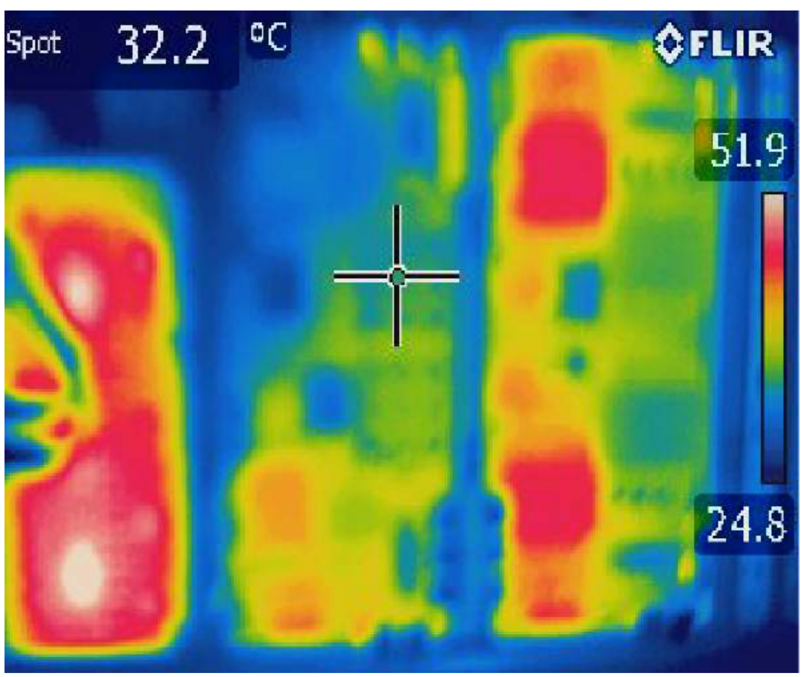
BETA publicatie	WP 526 (working paper)
ISBN	
ISSN	
NUR	
Eindhoven	March 2017

Maintenance Modeling and Optimization

Joachim Arts

Eindhoven University of Technology, School of Industrial Engineering, email: j.j.arts@tue.nl

March 3, 2017



Contents

Chapter 1

Introduction

> "A little neglect may breed great mischief...
> for want of a nail the shoe was lost
> for want of a shoe the horse was lost
> for want of a horse the rider was lost."
>
> Benjamin Franklin

The availability of capital goods is crucial to keep the primary processes of their owners/users up and running. Consider, aircraft, trains, wafer-steppers, and MRI scanners as examples. The inconvenience of trains and/or aircraft not running when needed (and planned) is a great inconvenience to travellers, but also a significant loss of revenue for airlines and railway operators. Wafer-steppers are used in the bottleneck production step of semi-conductor manufacturing. When a wafer-stepper is down, it causes the standstill of an entire semi-conductor factory. For the case of ASML wafer-steppers, these costs are in the order of magnitude of 100000 Euro per hour. The unavailibility of MRI-scanners is perhaps the most costly as it can lead to the loss of human life under some circumstances. All these examples illustrate that keeping capital assets up and running is of critical importance. Unfortunately, keeping capital assets up and running is also a costly business.

The costs of maintenance and unavailability of a capital asset over its lifetime (typically one to several decades) is a multiple of the acquisition price. Öner et al. (2007) estimate that these costs are 3 to 4 times the acquisition price for a specific engineer to order system. Van Dongen (2011) give several other examples where the maintenance costs alone are a multiple of the acquisition price as shown in Figure 1.1. Maintenance (including spare parts) and (unplanned) down-time of capital goods is one of the main culprits in these cost figures. In fact, in 2003 spare part sales

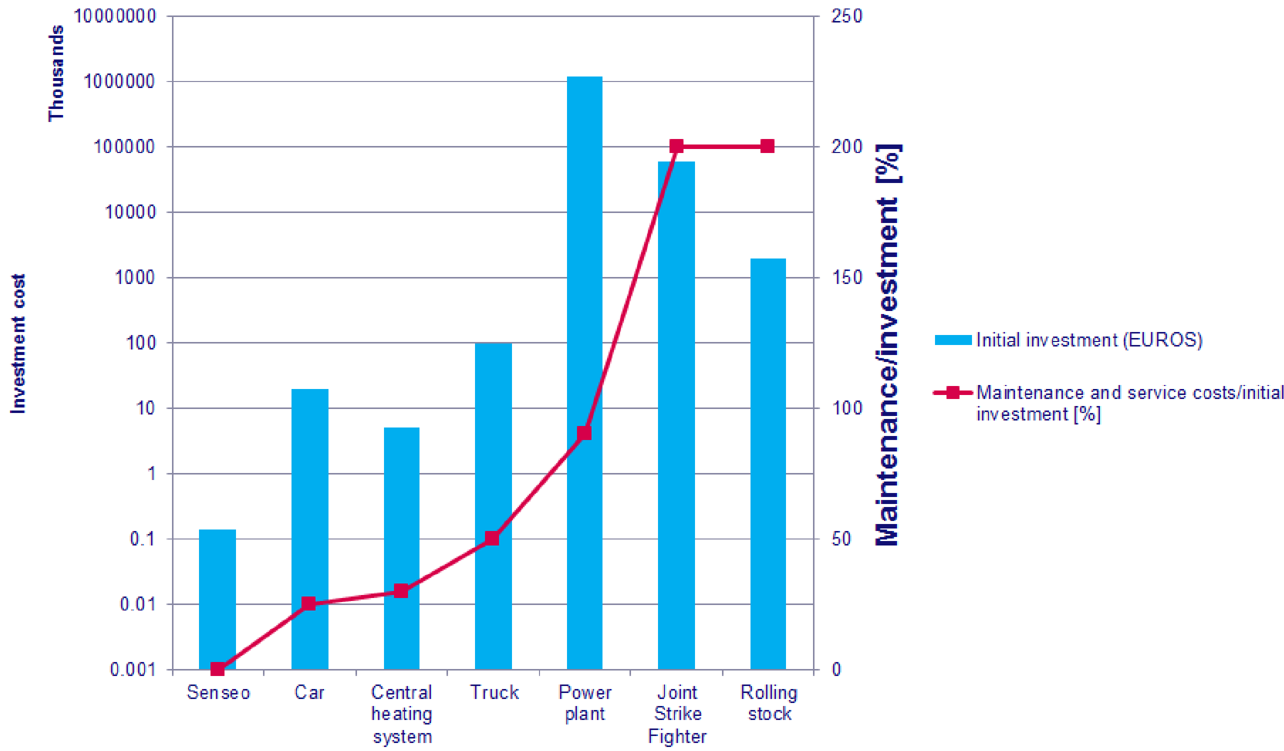


Figure 1.1: Acquisition costs versus maintenance costs over the lifetime for several assets

and services (mostly maintenance) accounted for 8% of the gross domestic product in the United States (AberdeenGroup, 2003). More recently, US bancorp estimated that the yearly expenditure in the US on spare parts amounts to 700 billion dollars which is 8% of the US gross domestic product (Jasper, 2006).

In these lecture notes, we describe several models to optimize maintenance operations, so let us first briefly consider maintenance operations. Different from regular production operations, maintenance operations are not instigated by demand from an outside customer, but by the need for maintenance of equipment. To perform maintenance, typically several resources are needed, the most important of which are:

- a specialist, mechanic, engineer or other trained professional
- tools and equipment
- spare parts.

In §1.1, we discuss different maintenance strategies and how they instigate the need for maintenance operations (and therefore also the resources mentioned above). The planning difficulties that arise in maintenance operations are discussed in §1.2.

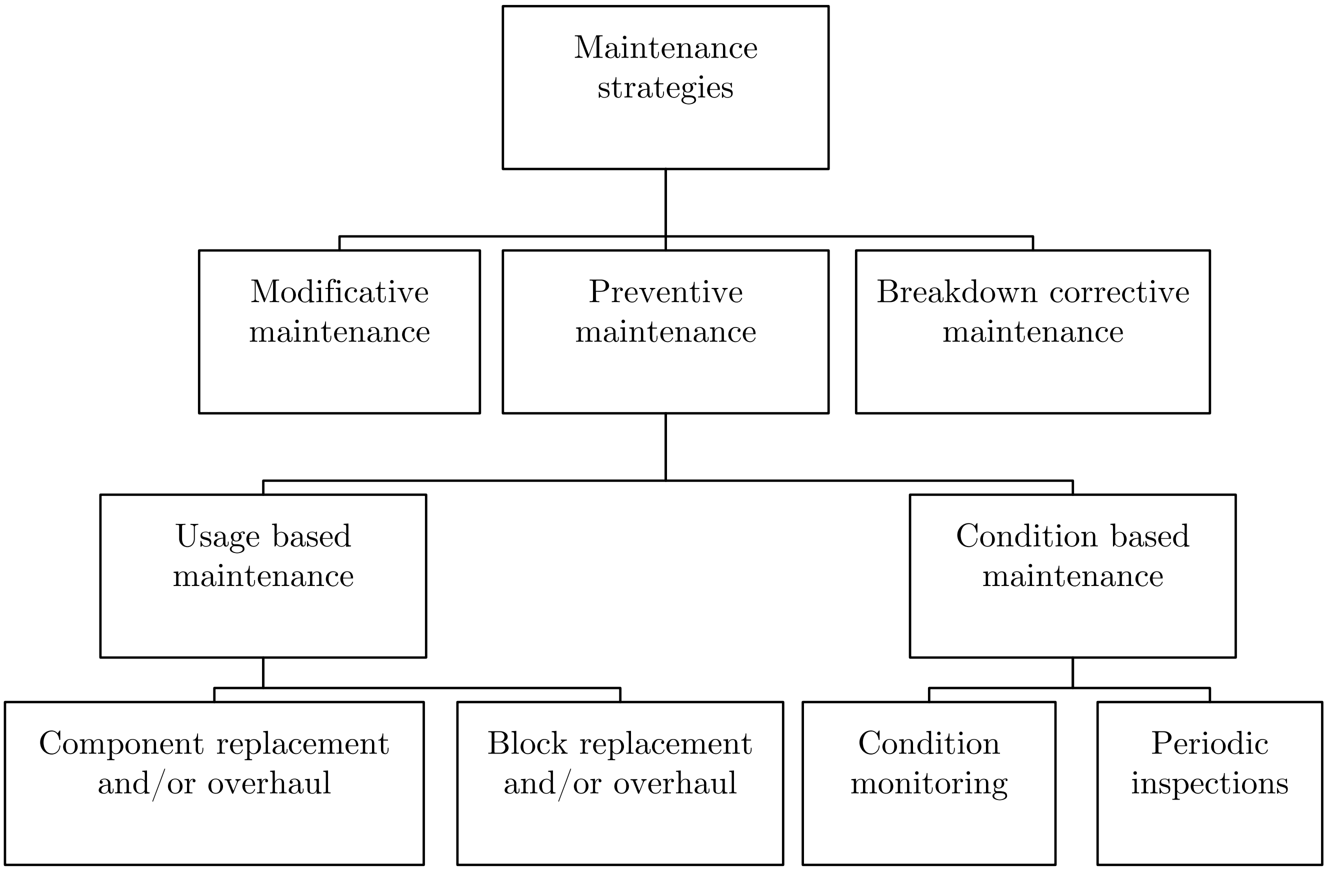


Figure 1.2: Maintenance strategies for components

1.1. Maintenance strategies

For the purpose of describing maintenance operations, it is convenient to think of equipment as a collection of interrelated components. Maintenance operations consist largely in replacing parts of equipment. Maintenance strategies determine when parts or equipment need to be replaced or maintained. Throughout this subsection, we focus on the decision to maintain/replace a component, but our discussion also applies to the decision to maintain/replace equipment. Figure 1.2[1] gives an overview of maintenance strategies. In this subsection, we follow Figure 1.2 in discussing different maintenance strategies.

Modificative maintenance concerns interchanging a part with a technically more advanced part in order to make the equipment perform better[2]. This form of maintenance is usually project

[1]Figure 1.2 was inspired by Figure 4.1 of Coetzee (1997), but has been significantly altered by the author.

[2]Sometimes maintenance is defined as any action that restores equipment to some previous state. Under this definition, modificative maintenance is an oxymoron.

based and non-recurring. The maintenance strategies that occur most often are preventive and breakdown corrective maintenance. Under a breakdown corrective maintenance strategy, a part is not replaced until it has failed, while under a preventive maintenance strategy, the aim is to replace parts before failure occurs. (Off course, this aim may not always be achieved: A part can break down before its planned preventive replacement occurs.) Breakdown corrective maintenance is an attractive option for components that do not wear, such as electronics. For parts that do wear, it can be beneficial to follow a preventive maintenance strategy.

Preventive maintenance strategies can be further divided into usage and condition based maintenance. Under usage based maintenance, the total usage of a part is measured and maintenance is conducted when a certain threshold level has been reached. The usage of parts can be measured in many ways depending on the nature of the equipment. Time in the field is perhaps the most common mean to measure usage. In these cases, usage based maintenance is also called age based maintenance. For vehicles (e.g., rolling stock), mileage is a common measure of usage. The number of on-off cycles is a measure of usage for equipment that is mainly loaded at the end or beginning of on-off cycles. For example, the number of landings is a measure of usage for the landing gear of an aircraft. Since the usage of equipment is usually scheduled, the moment that maintenance is performed can also be scheduled. If there is a large set-up cost associated with maintenance, it can be beneficial to interchange several parts simultaneously (Block replacement and/or overhaul). Otherwise, maintenance can be performed on a single component (Component replacement and/or overhaul).

In condition based maintenance, the actual condition of a part is gauged and maintenance is conducted based on this. The condition of a part can be measured either periodically during inspections (Periodic inspections) or continuously through a sensor (Condition monitoring). How the condition of equipment is measured depends on the nature of equipment. Below are some examples of how the condition of equipment can be measured:

- The condition of ball-bearings can be measured via the amplitude of vibrations around the bearing (Elwany and Gebraeel, 2008).

- The condition of a metal part can be determined by visually inspecting the number and length of cracks.

- For metal systems with moving parts, the concentration of ferrous parts in the lubrication fluid is measured as an indication of the wear and need for lubrication.

- The condition of a car engine is monitored continuously while driving by the engine-oil temperature gauge.

The need for maintenance can be ascertained periodically during an inspection or at any time in case of condition monitoring.

Which types of maintenance are prevalent for a given component depend very much on the technical nature of the equipment involved. Breakdown corrective maintenance is prevalent for electronic components that can be found in most high-tech assets. For aircraft, rolling stock and other heavy machinery with moving parts, the prevalent maintenance strategies are preventive (both usage and condition based).

1.2. Uncertainty in maintenance operations

Maintenance operations are subject to considerable uncertainty. For a particular maintenance action there can be uncertainty both with respect to its timing (When will we perform a maintenance action?) and content (What maintenance will we do during a maintenance action?). The different maintenance strategies discussed in the previous subsection are organized according to these two uncertainty dimensions in Table 1.1[3]

Table 1.1: Maintenance strategies organized by timing and content uncertainty.

		Timing of maintenance action	
		known	unknown
Content	known	Usage based or modificative maintenance	Condition based maintenance (Condition monitoring)
	unknown	Condition based maintenance (Periodic inspections)	Breakdown corrective maintenance

Usage based and modificative maintenance can be planned for ahead of time, whereas breakdown corrective maintenance cannot be planned for at all. As a consequence of this, the resources needed for usage based and modificative maintenance can be utilized more fully than resources needed for breakdown corrective maintenance.

Condition based maintenance is a hybrid form, in which some but not all uncertainty is taken away relative to breakdown corrective maintenance. Periodic inspections can be planned, and if they lead to maintenance, you know when the maintenance needs to be conducted (right after the

[3] Table 1.1 has been inspired by the maintenance box of Stoneham (1998) but has been altered significantly by the author.

inspection). However, the content of the maintenance depends on what is found during the inspection. Under condition monitoring, sensors provide realtime information about the degradation of equipment. The parts that need replacement can then be inferred from the sensor signal. However, degradation usually remains an uncertain process, so that the exact time that maintenance is needed remains unknown. Furthermore, there may be a time lapse between the observation of a signal and the opportunity to do maintenance. For vehicles (rolling stock, aircraft) we may need to wait until the vehicle is close to a maintenance depot. On the other hand, for production equipment or a windfarm at sea, we may need to wait until their operation can be stopped.

Remark 1.1. Sometimes the distinction between preventive and corrective maintenance is interpreted as being synonymous to planned and unplanned maintenance. This oversimplification only captures the upper left and lower right boxes of Table 1.1. Condition based maintenance is a hybrid form between planned and unplanned maintenance that deserves separate attention. $\diamond$

1.3. Structure of these lecture notes

The lecture notes are organized as follows. In §2, we introduce/review some elementary results from probability theory, reliability theory, and renewal processes. Readers familiar with this background can easily skip this chapter. Chapter 3 tackles several usage based maintenance optimization problems. All materials in this chapter are classical results except for the periodic age based policy with minimal repair in §3.2.5; these results are new and obtained especially for these lecture notes. Chapter 5 covers models for the condition and degradation of assets and/or components. These models are used in Chapter 5 to optimize decisions concerning maintenance based on the condition/degradation level. In the final chapter (6), we show how all the models in these lecture notes can be combined to design a complete maintenance concept/program for equipment.

Chapter 2

Short introduction/refresher in reliability theory, and renewal processes

"All knowledge degenerates into probability"

David Hume

In these lecture notes, we will mainly deal with non-negative random variables. Let the random variable T denote the *time to failure* of some component, $T \geq 0$. If T is a continuous random variable, we denote its distribution by $F_T(t) = \mathbb{P}(T \leq t)$ and assume it has a density $f_T(t) = \frac{d}{dt}F_T(t)$. The reliability of the component at time t is the probability the component survives beyond time t and is denoted by $R(t)$.

$$\begin{aligned} R(t) &= \mathbb{P}(T \geq t) \\ &= 1 - F_T(t) \\ &= 1 - \int_0^t f_T(t')dt' \\ &= \int_t^\infty f_T(t')dt'. \end{aligned} \tag{2.1}$$

The mean time to failure ($MTTF$) is just the expectation of T:

$$\begin{aligned}\mathbb{E}[T] &= \int t dF_T(t) \\ &= \int_0^\infty t f_T(t) dt \\ &= -\int_0^\infty t \frac{dR(t)}{dt} dt \\ &= [-tR(t)]_0^\infty + \int_0^\infty R(t) dt \\ &= \int_0^\infty R(t) dt. \end{aligned} \tag{2.2}$$

The third equality in (2.2) holds because $f_T(t) = \frac{dF_T(t)}{dt} = -\frac{dR(t)}{dt}$, the fourth equality follows from integration by parts and the final equality holds because $\lim_{t\to\infty} R(t) = 0$. The variance of T is given by:

$$\begin{aligned}\mathbf{Var}[T] &= \mathbb{E}\left[(T - \mathbb{E}[T])^2\right] \\ &= \mathbb{E}[T^2 - 2T\mathbb{E}[T] + (\mathbb{E}[T])^2] \\ &= \mathbb{E}[T^2] - 2\mathbb{E}[T]\mathbb{E}[T] - (\mathbb{E}[T])^2 \\ &= \mathbb{E}[T^2] - (\mathbb{E}[T])^2, \end{aligned} \tag{2.3}$$

where $\mathbb{E}[T^2] = \int t^2 dF_T(t) = \int_0^\infty t^2 f_T(t) dt$. The standard deviation of T is denoted by $\sigma_T = \sqrt{\mathbf{Var}[T]}$. Another measure of variation commonly used is referred to as the coefficient of variation, c_T, that satisfies:

$$c_T = \frac{\sigma_T}{\mathbb{E}[T]}. \tag{2.4}$$

Discrete random variables occur naturally in reliability engineering when a system degrades not with time but with the number of on-off cycles. An example of this was given in the introduction: Landing gear of aircraft degrade with the number of landings, not with time or mileage. For all the results in this section, there are straightforward equivalents for discrete random variables. Integrals are replaced by summations in these results.

2.1. Failure rates

The distribution function and the density function of the time to failure provide only limited understanding of the physical failure mechanisms that cause failures. To gain some understanding, let us consider a time t at which the component has not failed yet. We ask ourselves the following question: How likely is this component to fail in the next (small) time interval of length ε relative to the length of this interval? More formally we would like to know:

$$h(t) = \lim_{\varepsilon \downarrow 0} \frac{\mathbb{P}(T \leq t + \varepsilon | T \geq t)}{\varepsilon}. \tag{2.5}$$

for t such that $f_T(t) > 0$. This function can be loosely interpreted as the instantaneous expected number of failures per time unit at time t, and is known as the *failure rate*, *hazard rate* (hence the notation $h(t)$) or *mortality rate*[1]. We will refer to it as the failure rate. Now using basic probability we can write for the probability in (2.5)

$$\begin{aligned} \mathbb{P}(T \leq t + \varepsilon | T \geq t) &= \frac{\mathbb{P}(T \leq t + \varepsilon \cap T \geq t)}{\mathbb{P}(T \geq t)} \\ &= \frac{\mathbb{P}(t \leq T \leq t + \varepsilon)}{\mathbb{P}(T \geq t)} \\ &= \frac{F_T(t+\varepsilon) - F_T(t)}{R(t)}. \end{aligned} \tag{2.6}$$

Reinserting (2.6) into (2.5) we find

$$h(t) = \lim_{\varepsilon \downarrow 0} \frac{F_T(t+\varepsilon) - F_T(t)}{\varepsilon} \frac{1}{R(t)} = \frac{f_T(t)}{R(t)}, \tag{2.7}$$

where we used the definition of a derivative and the fact that $f_T(t)$ is the derivative of $F_T(t)$. The identity in (2.7) is the definition of the failure rate most often used in textbooks.

The failure rate reveals some essential features about the degradation process of components. If the failure rate is an increasing function of time $\left(\frac{dh(t)}{dt} > 0\right)$, the component degrades over time. If this is the case, we say that the time to failure is *IFR*. (IFR is the abbreviation for increasing failure rate.) Mechanical devices typically have an increasing failure rate.

If the failure rate is a decreasing function of time $\left(\frac{dh(t)}{dt} < 0\right)$, the component becomes more reliable over time (conditional on not having failed already). If this is the case, we say that the time to failure is *DFR*. (DFR, as you might have guessed, is the abbreviation for decreasing failure rate.) Electronic components often have DFR. The reason for this is that electronics are not usually subject to wear (in clean conditions at least) unless there is a manufacturing defect. However, the longer an electronic components has been functioning without problems, the more likely it is that there is no manufacturing defect.

A special case that we will consider separately in §2.2 is where the failure rate is constant (*CFR*), $\left(\frac{dh(t)}{dt} = 0\right)$. (An even more peculiar case occurs when the failure rate alternates from

[1] The statistical development of reliability engineering closely follows the statistical development of studies in the medical and life sciences. Some of their terminology is still used in our field, even though equipment does not die, patients or lab-rats do.

being decreasing and increasing over time like a sine function, say. In this course, we will not use such "exotic" distributions.)

There is also a discrete equivalent to the failure rate. If T is a discrete random variable on the (non-negative) integers, let $p_k = \mathbb{P}(T = k)$ and $R_k = \mathbb{P}(T \geq k)$. The discrete equivalent of the failure rate, h_k, is then the probability of the component failing after k cycles conditional on it surviving at least k cycles:

$$h_k = \mathbb{P}(T = k | T \geq k) = \frac{\mathbb{P}(T = k \cap T \geq k)}{\mathbb{P}(T \geq k)} = \frac{p_k}{R_k}. \tag{2.8}$$

Note that $0 \leq h_k \leq 1$. The discrete hazard rate is perhaps slightly more intuitive because its definition does not involve limits and other devices from calculus.

We close this section by showing how the density and reliability function can be obtained from the failure rate. First observe that $f_T(t) = \frac{d}{dt}F_T(t) = \frac{d}{dt}(1 - R(t)) = -\frac{d}{dt}R(t)$. Inserting this into (2.7) yields:

$$h(t) = -\frac{1}{R(t)}\frac{dR(t)}{dt}. \tag{2.9}$$

Now taking (2.9), multiplying both sides by dt and integrating between 0 and t yields

$$\int_0^t h(u)du = -\int_0^t \frac{1}{R(t)}dR(t) = -\ln(R(t)). \tag{2.10}$$

Finally, solving (2.10) for $R(t)$ we find:

$$R(t) = \exp\left(-\int_0^t h(u)du\right). \tag{2.11}$$

2.2. Commonly used distributions

In this subsection, we give some results on distributions used in maintenance and reliability engineering. Before doing this, we cover some groundwork. The factorial is defined for all positive integer numbers, e.g., $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$. Sometimes it is convenient to work with an extension of the factorial so that it also applies to real numbers. The Gamma-function achieves this and is defined as follows:

$$\Gamma(x) = \int_0^\infty \exp(-u)u^{x-1}du, \quad x > 0 \tag{2.12}$$

The Gamma-function has the following properties:

- $\Gamma(x + 1) = x\Gamma(x), \quad x > 0$

- $\Gamma(1) = 1$
- $\Gamma(x) = (x-1)!, \quad x = 1, 2, \cdots$

For values between 1 and 3, the Gamma-function is evaluated in Table 2.1. Using the first property of the Gamma-function, this table can be used to evaluate the Gamma-function for larger real numbers. For example:

$$\Gamma(5.6) = 4.6\Gamma(4.6) = 4.6 \cdot 3.6 \cdot \Gamma(3.6) = 4.6 \cdot 3.6 \cdot 2.6 \cdot \Gamma(2.6) = 4.6 \cdot 3.6 \cdot 2.6 \cdot 1.4296 = 61.55.$$

Table 2.1: The Gamma-function for arguments between 1 and 3.

	$\Gamma(x+y)$									
$x \backslash y$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.0000	0.9514	0.9182	0.8975	0.8873	0.8862	0.8935	0.9086	0.9314	0.9618
2	1.0000	1.0465	1.1018	1.1667	1.2422	1.3293	1.4296	1.5447	1.6765	1.8274

In §2.2.1-§2.2.5, we discuss continuous distribution functions and in §2.2.6-§2.2.9, we discuss discrete distribution functions.

2.2.1 Exponential distribution

An exponential random variable X has the following density and distribution function:

$$f(x) = \lambda \exp(-\lambda x), \qquad F(x) = 1 - \exp(-\lambda x), \qquad x \geq 0$$

which is supported on $[0, \infty)$. The parameter $\lambda > 0$ is called the scale parameter. The mean, variance and coefficient of variation of X are given by:

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \mathbf{Var}[X] = \frac{1}{\lambda^2}, \quad c_X = \sqrt{\frac{\mathbf{Var}[X]}{\mathbb{E}^2[X]}} = 1.$$

The exponential distribution is important in operations management and reliability engineering because it has the *lack of memory* property. This property means that the remaining lifetime of a part has the same distribution as the original lifetime:

$$\begin{aligned}\mathbb{P}(X > s+t | X > s) &= \mathbb{P}(X > s+t \cap X > s)/\mathbb{P}(X > s)\\ &= \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > s)}\\ &= \frac{\exp(-\lambda(s+t))}{\exp(-\lambda s)}\\ &= \exp(-\lambda t)\\ &= \mathbb{P}(X > t).\end{aligned}$$

Another special property of the exponential distribution, which is actually equivalent to the lack of memory property, is that it has a constant failure rate

$$h(t) = \lambda.$$

This means that the exponential distribution is the unique distribution that is both IFR and DFR.

2.2.2 Uniform distribution

The uniform distribution is supported on (a, b) and has density and distribution

$$f(x) = \begin{cases} 1/(b-a), & a \leq x \leq b;\\ 0, & \text{otherwise.}\end{cases} \qquad F(x) = \begin{cases} (x-a)/(b-a), & a \leq x \leq b;\\ 0, & x \leq a;\\ 1, & b \leq x.\end{cases}$$

The reason for the name uniform should be obvious from Figure 2.1.

If X has a uniform distribution on (a, b), then the mean, variance, and coefficient of variation are given by:

$$\mathbb{E}[X] = \frac{a+b}{2}, \quad \mathbf{Var}[X] = \frac{(b-a)^2}{12}, \quad c_X = \frac{b-a}{\sqrt{3}(a+b)} \tag{2.13}$$

The uniform distribution has an increasing failure rate:

$$h(x) = \begin{cases} 0, & x < a;\\ \frac{1}{b-x}, & a \leq x \leq b.\end{cases} \tag{2.14}$$

2.2.3 Erlang distribution

If $E_1, E_2, \cdots E_k$ are i.i.d. exponential random variables with mean λ^{-1}, then $X = \sum_{i=1}^{k} E_i$ has an Erlang[2] distribution with *shape parameter* $k \in \mathbb{N}$ and *scale parameter* $\lambda > 0$. Like the exponential

[2] The Erlang distribution has been named after the Danish engineer Agner Krarup Erlang (1878-1929). Erlang is considered one of the founders of queueing theory. The Erlang distribution is often used in this field.

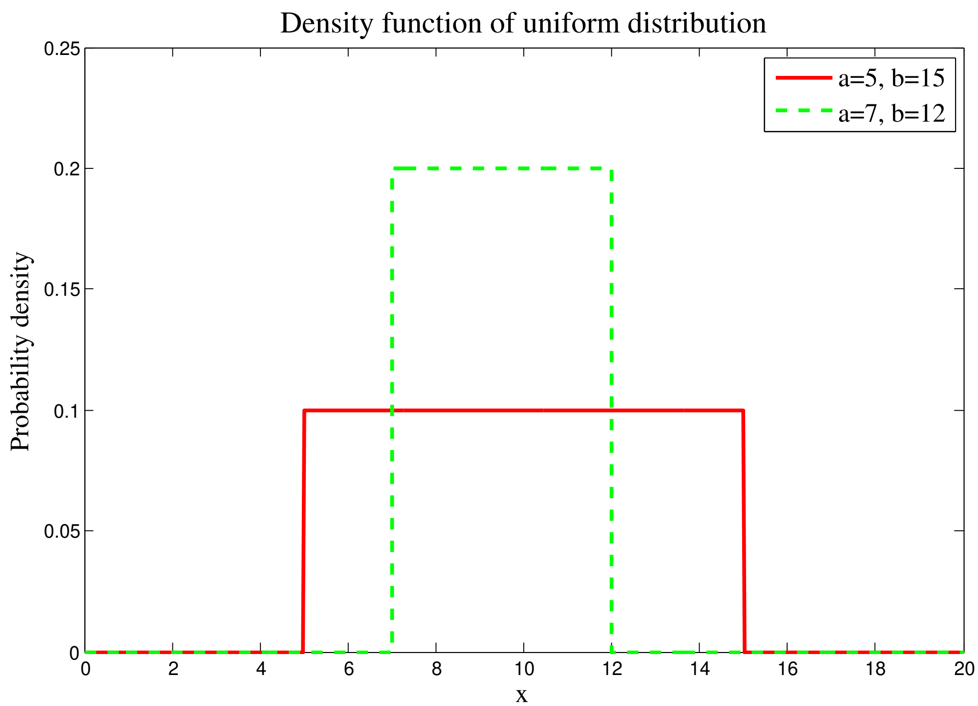


Figure 2.1: The probability density function of the Uniform distribution for several parameter values.

distribution, the Erlang distribution has support on $[0, \infty)$ and its density and distribution function are given by:

$$f(x) = \frac{\lambda^k x^{k-1}}{(k-1)!} \exp(-\lambda x), \quad F(x) = 1 - \sum_{n=0}^{k-1} \frac{(\lambda x)^n}{n!} \exp(-\lambda x), \quad x \geq 0. \tag{2.15}$$

The mean, variance and coefficient of variation of X are given by:

$$\mathbb{E}[X] = \frac{k}{\lambda}, \quad \mathbf{Var}[X] = \frac{k}{\lambda^2}, \quad c_X = \frac{1}{\sqrt{k}}. \tag{2.16}$$

The failure rate of an Erlang random variable is given by:

$$h(x) = \lambda \left[\sum_{n=0}^{k-1} \frac{(k-1)!}{n!(\lambda x)^{k-1-n}} \right]^{-1}. \tag{2.17}$$

The failure rate of the Erlang distribution is constant for $k = 1$. (In fact, it reduces to the exponential distribution when $k = 1$.) For $k > 1$, the Erlang distribution has an increasing failure rate. Figure 2.2 shows the Erlang density function for several shape and scale parameters.

Finally, we note from the construction of the Erlang distribution that if Y_1 has an Erlang distribution with shape k_1 and scale λ and Y_2 has en Erlang distribution with shape k_2 and scale λ, then $Z = Y_1 + Y_2$ has an Erlang distribution with shape $k_1 + k_2$ and scale λ.

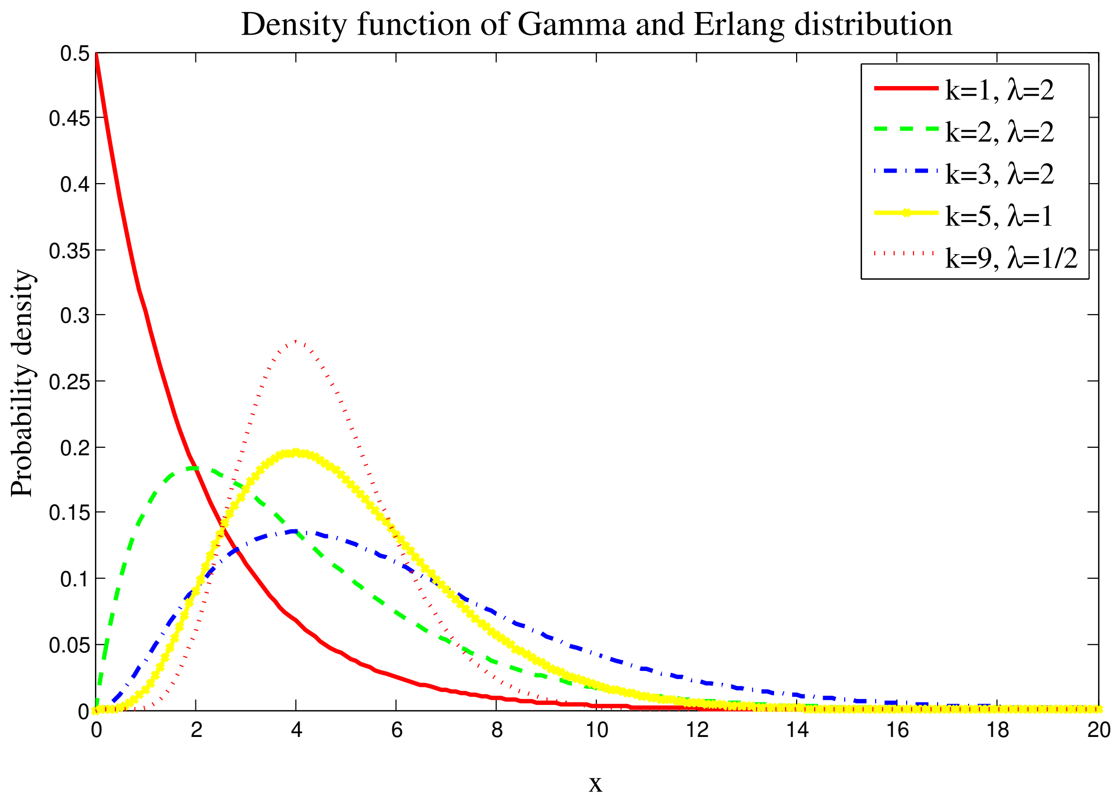


Figure 2.2: The Erlang (and Gamma) probability density function for several shape and scale parameters.

2.2.4 Gamma distribution

The Erlang distribution can be generalized by allowing k to take non-integer values. In this case, all the factorials in (2.15) and (2.17) have to be replaced by their equivalents in terms of the Gamma-function. This distribution is called the Gamma distribution and is parameterized by the *shape parameter* $\alpha > 0$ (which is equivalent to k in the Erlang distribution), and the *scale parameter* $\beta > 0$ (which is equivalent to λ in the Erlang distribution). Its density and distribution are given by:

$$f(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta x), \quad F(x) = \int_0^x f(u)du, \quad x \geq 0. \tag{2.18}$$

The mean, variance and coefficient of variation of a Gamma distributed random variable X are given by:

$$\mathbb{E}[X] = \frac{\alpha}{\beta}, \quad \mathbf{Var}[X] = \frac{\alpha}{\beta^2}, \quad c_X = \frac{1}{\sqrt{\alpha}} \tag{2.19}$$

The failure rate of a Gamma distribution is given by:

$$h(x) = \left[\int_x^\infty \left(\frac{u}{x}\right)^{\alpha-1} \exp(-\beta(u-x))du\right]^{-1} = \left[\int_0^\infty (1+v/x)^{\alpha-1} \exp(-\beta v)dv\right]^{-1}, \quad x \geq 0. \tag{2.20}$$

For $0 < \alpha < 1$, this failure rate is decreasing; for $\alpha > 1$ this failure rate is increasing and for $\alpha = 1$, we obtain the constant failure rate.

Gamma random variables also inherit the following property from the Erlang distribution: If Y_1 has a Gamma distribution with shape α_1 and scale β, and Y_2 has a Gamma distribution with shape α_2 and scale β, then $Z = Y_1 + Y_2$ has a Gamma distribution with shape $\alpha_1 + \alpha_2$ and scale β.

2.2.5 Weibul distribution

The Weibul[3] distribution is used much in reliability engineering because it provides a good fit with data in many applications and arises naturally in theory. Many components are composed of even smaller subcomponents. The time until failure of the component is therefore the shortest time until failure of any of the subcomponents. If we let $Y_1, \cdots, Y_n$ denote the times until failure of the subcomponents, the time until failure of the component itself is $X = \min\{Y_1, Y_2, \cdots, Y_n\}$. If the random variables Y_1 to Y_n have finite support, then the distribution of X will approach the Weibul distribution as n approaches infinity[4]. This is the reason why the Weibul distribution arises theoretically and often provides good empirical fit with data.

The Weibul distribution has *shape* parameter β and *scale* parameter η. Its density and distribution are:

$$f(x) = \frac{\beta x^{\beta-1}}{\eta^\beta} \exp\left(-(x/\eta)^\beta\right), \quad F(x) = 1 - \exp\left(-(x/\eta)^\beta\right), \quad x \geq 0. \tag{2.21}$$

Note that for $\beta = 1$, the Weibul distribution reduces to the exponential distribution. The mean and variance are given by:

$$\mathbb{E}[X] = \Gamma(1+1/\beta)\eta, \quad \mathbf{Var}[X] = \left[\Gamma(1+2/\beta) - (\Gamma(1+1/\beta))^2\right]\eta^2. \tag{2.22}$$

[3]The Weibul distribution has been named after Ernst Hjalmar Waloddi Weibull (1887-1979) who was a Swedish engineer. He did not invent the Weibul distribution, but he made it popular among (reliability) engineers, because he showed that it arises naturally in the study of strength of materials, fatigue, rupture in solids and bearings.

[4]Compare this result to the central limit theorem. The central limit theorem plays an important role in statistics and says that $X = Y_1 + Y_2 + \cdots Y_n$ approaches a normal random variables as $n \to \infty$ (if the Y_i are independent and have finite first two moments). In reliability engineering, we are more often interested in $X = \min\{Y_1, Y_2, \cdots, Y_n\}$ which converges to a Weibul random variable as $n \to \infty$ (if the Y_i are non-negative and have finite support.) This result is known as the type III extreme value law.

The failure rate of the Weibul distribution has an exponential form:

$$h(x) = \frac{\beta}{\eta^\beta} x^{\beta-1}. \tag{2.23}$$

From this form, we immediately observe that X is IFR for $\beta > 1$, DFR for $0 < \beta < 1$, and CFR for $\beta = 1$. Figure 2.3 shows the Weibul probability density function for several different shape parameters.

We finish this section by highlighting an important property of Weibul random variables. Suppose that $X_1, X_2, \ldots, X_n$ are independently identically Weibul distributed random variables with shape parameter β and scale parameter η. Think of X_1 to X_n as the lifetimes of n components in a serial configuration. Let us call the time to failure of this system Y; it obviously satisfies $Y = \min(X_1, X_2, \ldots, X_n)$. Now let us consider the reliability of the entire system:

$$\begin{aligned}
\mathbb{P}(Y > y) &= \mathbb{P}(X_1 > y \cap X_2 \geq y \cap \ldots \cap X_n \geq y) \\
&= \mathbb{P}(X_1 > y)\mathbb{P}(X_2 > Y) \cdots \mathbb{P}(X_n > y) \\
&= \left[\exp\left(-\left(\frac{y}{\eta}\right)^\beta\right)\right]^n \\
&= \exp\left(-n\left(\frac{y}{\eta}\right)^\beta\right) \\
&= \exp\left(-\left(\frac{y}{\eta/n^{1/\beta}}\right)^\beta\right).
\end{aligned} \tag{2.24}$$

We have thus shown that Y also has a Weibul distribution with shape parameter $\beta' = \beta$ and scale parameter $\eta' = \eta n^{-1/\beta}$. This result gives some intuition on why the Weibul distribution is such a prevalent model in reliability engineering.

2.2.6 Poisson distribution

The Poisson[5] distribution is a discrete distribution on the non-negative integers characterized by its mean μ only. If the random variable X has Poisson distribution with mean μ then

$$p_x = \mathbb{P}(X = x) = \exp(-\mu)\frac{\mu^x}{x!}, \quad x \in \mathbb{N}_0 = \mathbb{N} \cap \{0\} \tag{2.25}$$

and

$$\mathbb{E}[X] = \mu, \quad \mathbf{Var}[X] = \mu, \quad c_X = \frac{1}{\sqrt{\mu}}. \tag{2.26}$$

Poisson probabilities satisfy the following recursive relation that is convenient in computations:

$$p_x = \frac{\mu}{x} p_{x-1}, \quad x \in \mathbb{N}. \tag{2.27}$$

[5]The Poisson distribution is called after the French mathematician Siméon Poisson (1781-1840).

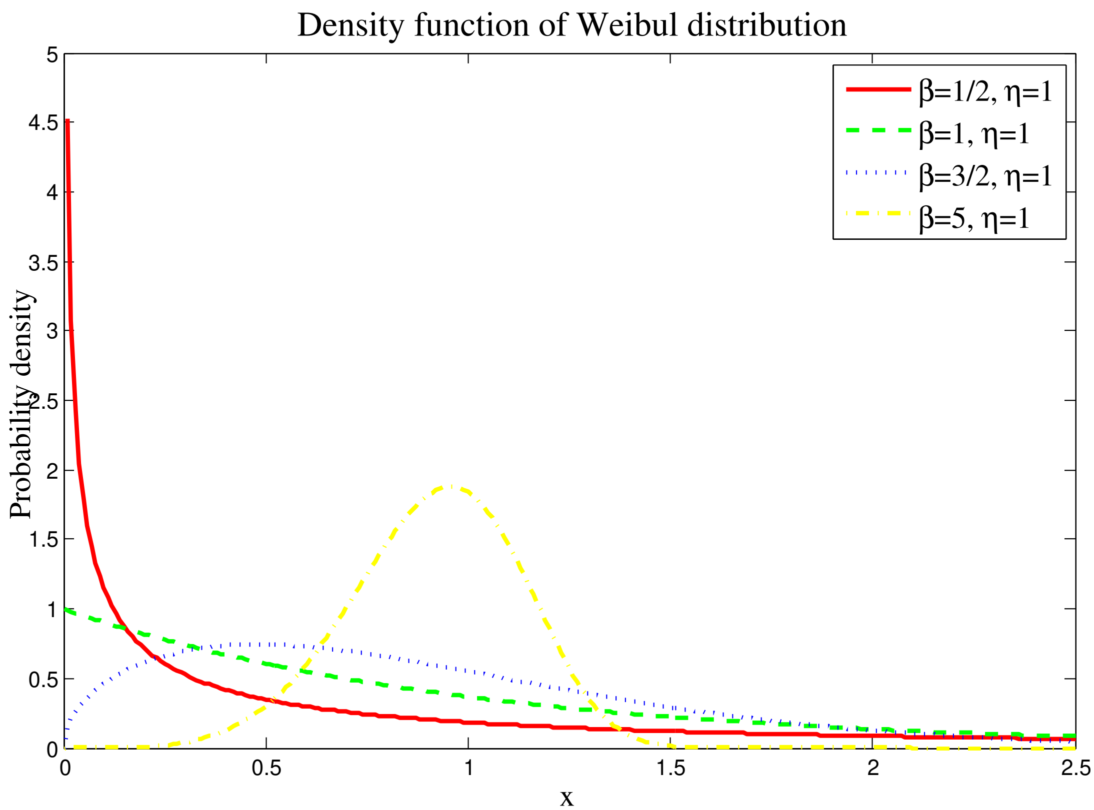


Figure 2.3: The Weibul probability density function for several shape parameters

2.2.7 Geometric distribution

The geometric distribution is a discrete distribution characterized by a single parameter $p \in (0, 1)$ that represents the failure probability. Suppose that we periodically load a system and there is probability p that a system fails during any loading. If we let X denote the number of times the system can be loaded *without* failure, then X is said to have a geometric distribution with failure probability p[6]. The geometric distribution is memoryless and can be interpreted as the discrete analogue of the exponential distribution. Its probability mass and distribution functions are given by

$$p_x = \mathbb{P}(X = x) = p(1-p)^x, \quad \mathbb{P}(X \leq x) = 1 - (1-p)^x, \quad x \in \mathbb{N}_0. \tag{2.28}$$

The mean and variance are given by

$$\mathbb{E}[X] = \frac{1-p}{p}, \quad \mathbf{Var}[X] = \frac{1-p}{p^2}. \tag{2.29}$$

[6]The geometric distribution is also used to model the number events until a successful event occurs; think for example of the number of coin tosses until heads comes up. The convention to call p the failure probability is context dependent. In other contexts it can be referred to as the success-probability.

The failure rate of the geometric distribution is constant and equal to p. (This follows immediately from the interpretation of the geometric distribution.)

2.2.8 Negative binomial distribution

If $G_1, G_2, \dots, G_r$ are i.i.d. geometric random variables with failure probability p, then $X = \sum_{i=1}^{r} G_i$ has a negative binomial distribution with *shape* parameter r and *scale* parameter p. We have:

$$p_x = \mathbb{P}(X = x) = \binom{r+x-1}{x} p^r (1-p)^x, \quad P(X \leq x) = \sum_{k=0}^{x} p_k, \quad x \in \mathbb{N}_0. \tag{2.30}$$

By its definition it is easy to find that

$$\mathbb{E}[X] = r\mathbb{E}[G_1] = r\frac{1-p}{p}, \quad \mathbf{Var}[X] = r\mathbf{Var}[G_1] = r\frac{1-p}{p^2}. \tag{2.31}$$

Also note (again from definition) that if Y_1 has a negative binomial distribution with shape r_1 and scale p, and Y_2 has a negative binomial distribution with shape r_2 and scale p, then $Z = Y_1 + Y_2$ has a negative binomial distribution with shape $r_1 + r_2$ and scale p.

Just like the Gamma distribution generalizes the Erlang distribution by allowing the shape parameter to be non-integer, the negative binomial distribution can also be generalized by allowing non-integer r, but its name remains the same. Everything described above still holds in this case, but same care needs to be taken in interpreting the binomial coefficient $\binom{r+x-1}{x} = \frac{(r+x-1)!}{x!(r-1)!}$ so that it also works for non-integer r. This can be done as follows:

$$\begin{aligned}
\binom{r+x-1}{x} &= \frac{(r+x-1)!}{x!(r-1)!} \\
&= \frac{(r+x-1)(r+x-2)(r+x-3)\cdots(r-1)(r-2)(r-3)\cdots 1}{x(x-1)(x-2)(x-3)\cdots 1\cdot(r-1)(r-2)(r-3)\cdots 1} \\
&= \frac{(r+x-1)(r+x-2)(r+x-3)\cdots r}{x(x-1)(x-2)\cdots 1} \\
&= \prod_{k=1}^{x} \frac{r+x-k}{k}.
\end{aligned} \tag{2.32}$$

The last expression can also be computed for non-integer r and should be used also for integer r because it is numerically more stable. The most convenient way to compute the pmf is by using the recursive relation

$$\mathbb{P}(X = x) = \frac{r+x-1}{x}(1-p)\mathbb{P}(X = x-1), \quad P(X = 0) = p^r, \quad x \in \mathbb{N}. \tag{2.33}$$

2.2.9 Logarithmic distribution

If X has a logarithmic distribution with parameter $q \in (0,1)$, then

$$p_x = \mathbb{P}(X = x) = \frac{-1}{\ln(1-q)} \frac{q^x}{x}, \quad \mathbb{P}(X \leq x) = \sum_{i=0}^{x} p_i, \quad x \in \mathbb{N}, \tag{2.34}$$

and

$$\mathbb{E}[X] = \frac{-1}{\ln(1-q)} \frac{q}{1-q}, \quad \mathbf{Var}[X] = -q \frac{q + \ln(1-q)}{(1-q)^2(\ln(1-q))^2}. \tag{2.35}$$

Figure 2.4 shows the logarithmic probability mass function for several q. For computational purposes it is convenient that the probability mass function can be computed recursively as

$$p_{x+1} = \frac{q}{x+1} p_x. \tag{2.36}$$

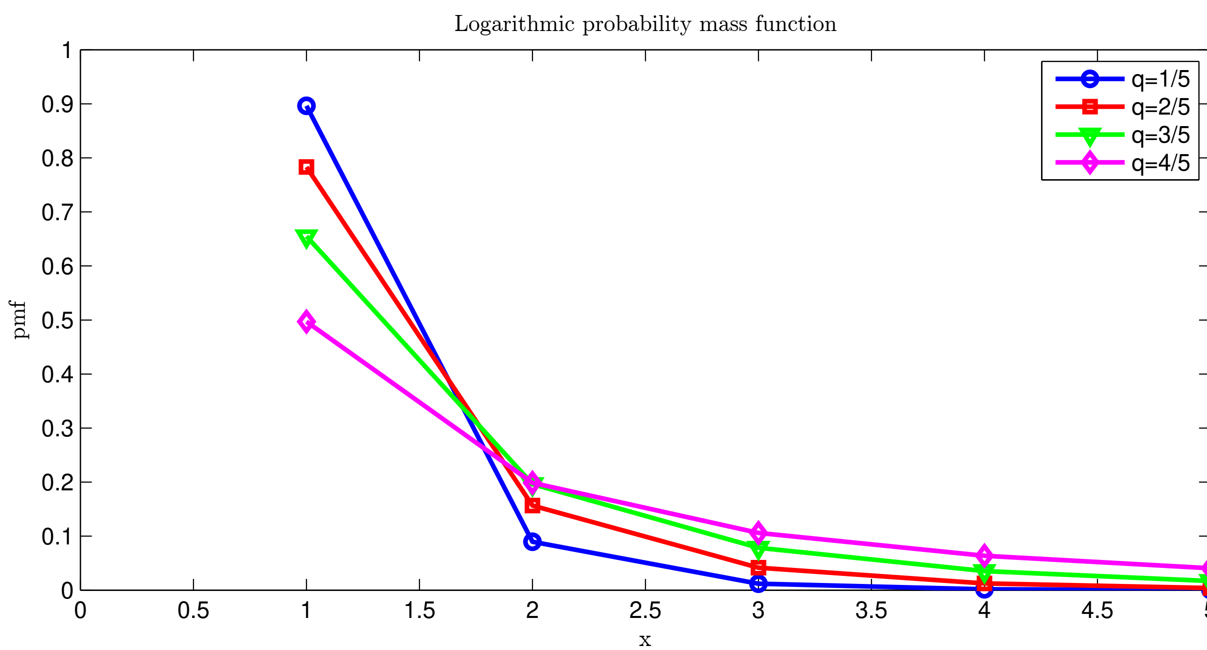


Figure 2.4: The Logarithmic probability mass function for several q

2.3. Renewal theory

A renewal process is a counting process in which the time between events (also called renewals) are independently and identically distributed (i.i.d.). Let $X_1, X_2, \cdots$ be a sequence of non-negative i.i.d. random variables with common distribution $F(x)$, density $f(x) = \frac{dF(x)}{dx}$, and mean $0 <$

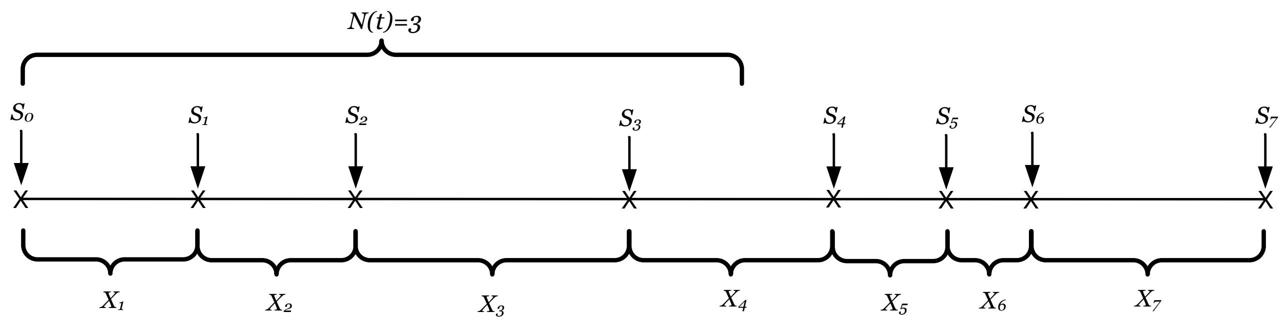


Figure 2.5: This Figure shows a sample path of a renewal process with the corresponding notation. Each renewal is marked by an x on the time line.

$\mathbb{E}[X_i] < \infty$. X_i is the time between the $(i-1)$-th and i-th renewal; for example X_i might represent the time to failure of a component. Now define the time the i-th renewal occurs as S_i:

$$S_i = \sum_{k=1}^{i} X_k, \qquad S_0 = 0. \tag{2.37}$$

If the X_i represent the time to failure of some component, then S_i represents the time until the failure of the i-th component. We may ask ourselves how many renewals have occurred up until time t. Such a process is called a renewal process and denoted by $N(t)$:

$$N(t) = \max\{i \in \mathbb{N}_0 | S_i \leq t\}, \quad t \geq 0. \tag{2.38}$$

Figure 2.5 shows an example of a renewal process with the notation that we introduced. The expected number of renewals up until time t, $\mathbb{E}[N(t)]$, is called the renewal function and denoted by $M(t)$. The renewal function obeys the following integral equation.

Theorem 2.1. (Renewal equation) *The renewal function $M(t) = \mathbb{E}[N(t)]$ satisfies:*

$$M(t) = F(t) + \int_0^t M(t-x)f(x)dx, \quad t \geq 0 \tag{2.39}$$

Proof. The proof follows from conditioning on X_1:

$$\begin{aligned} M(t) &= \mathbb{E}[N(t)] \\ &= \int_0^\infty \mathbb{E}[N(t)|X_1 = x]f(x)dx \\ &= \int_0^t \mathbb{E}[N(t)|X_1 = x]f(x)dx \\ &= \int_0^t (1 + M(t-x))f(x)dx \\ &= F(t) + \int_0^t M(t-x)f(x)dx \end{aligned}$$

The third equality follows because $N(t) = 0$ if $X_1 \geq t$. The fourth equality follows because if $X_1 = x \leq t$, then at least one renewal already occurred, and the expected number of renewals from x to t is just $M(t-x)$. □

Equation (2.39) is not immediately helpful for computing $M(t)$ because solving an integral equation (such as (2.39)) is, in general, more difficult than solving a differential equation. If you can guess the right form for $M(t)$ under some $f(x)$ and $F(x)$, you can plug it into (2.39) to check if this guess is correct. Coming up with such a guess is difficult and usually requires detailed understanding of the context.

There is also a discrete equivalent to Theorem 2.1, which is more immediately helpful in computations. Exercise 1.5 explores this further.

Example 2.1. The *Poisson process* is a special case of a renewal process for which $X_1, X_2, \cdots$ are exponentially distributed with mean λ^{-1}, S_k has an Erlang distribution with shape parameter k and scale parameter λ and $N(t)$ has a Poisson distribution with mean λt. From these results we immediately have that the renewal function satisfies $M(t) = \lambda t$. ◇

2.4. Renewal reward theory

Now suppose that there is a reward (cost), W_i, associated with each renewal i. Furthermore, assume that $W_1, W_2, \cdots$ is an iid sequence of random variables with $|\mathbb{E}[W_i]| < \infty$. (Note that we do *not* assume that W_i and X_i are independent.) The total reward (cost) up until time t is denoted by $Y(t)$ and satisfies:

$$Y(t) = \sum_{i=1}^{N(t)} W_i. \tag{2.40}$$

$Y(t)$ is called a renewal reward process. A natural question that arises is: "What is the average reward (costs) per time unit for a renewal reward process?". The following theorem addresses this. The proof of this theorem is rather technical so we omit it. The interested reader is referred to Ross (1996) Section 3.6 for a formal proof.

Theorem 2.2. (Renewal reward theorem) *The average reward (cost) per time unit of a renewal process, g, satisfies*

$$g = \lim_{t\to\infty} \frac{Y(t)}{t} = \lim_{t\to\infty} \frac{\mathbb{E}[Y(t)]}{t} = \frac{\mathbb{E}[W_i]}{\mathbb{E}[X_i]}$$

The expected length of a renewal is sometimes called the expected cycle length, ECL. The average costs per renewal is sometimes referred to as the expected cycle cost, ECC. Theorem 2.2, also known as the renewal reward theorem, says that that he expected cost per time unit is simply ECC/ECL.

Chapter 3

Usage Based Maintenance Models

> "An ounce of prevention is worth a pound of cure"
>
> Benjamin Franklin

In this chapter, we are primarily concerned with the questions: "When should we conduct maintenance and what are the maintenance costs per time unit?". The answers to these questions depend very much on the setting you consider. Section 3.1 considers setting in which degradation is deterministic and Section 3.2 considers several setting where the time until failure is stochastic.

3.1. Deterministic maintenance models

The time to failure was presented as clear cut moment in time (that may be random) in the preceding discussion. Many components do not have such a clear failure moment, but their performance degrades over time. Consider for example the rubber belt of a conveyor; see Figure 3.1. The rubber belt loses elasticity as a result of usage in a predictable manner. This elasticity loss causes slippage of the conveyor, which causes the production to slow down. In this situation, the production loss is a reason to replace the belt before it has actually failed. Other examples of components for which this holds are fuel filters, the felt of a paper mill, and molds in glass and plastic production. The usage based maintenance policy that is explained in the introduction (see also Figure 1.2) is an appropriate maintenance policy in this situation. We consider the policy in which a component is replaced after τ usage. (We will assume that usage is measured in time units, but this can be changed to mileage, number of produced products etc.) The question under

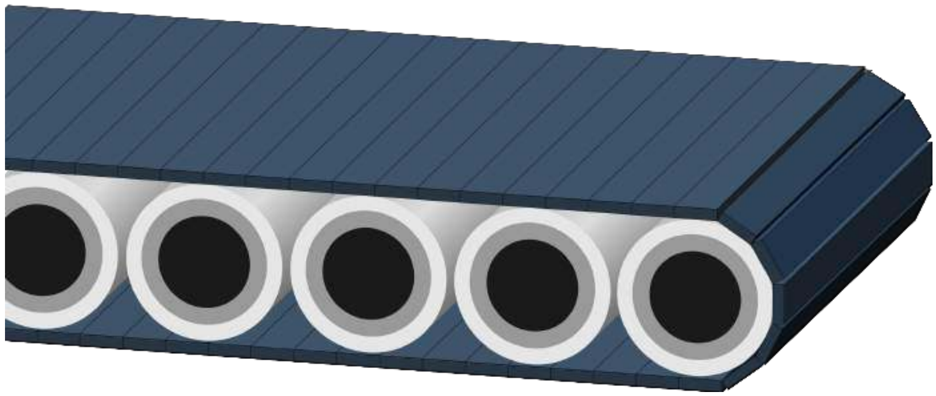


Figure 3.1: The actual rubber belt on a conveyor loses elasticity over time which causes slippage.

this policy is: "After what amount of usage should the component (belt) be replaced, i.e., what is the optimal τ?". This section provides a mathematical model to answer that question.

Suppose that the costs for replacing the belt are $C_p > 0$ euros. (The subscript p indicates the replacement is planned and the same notation will be used later.) Over usage time t, the belt slows down which causes the production rate (measured in euros per time unit) to decrease by $a > 0$ euros per time unit per time unit. (So if $a = 5$ euro/hr^2, and the production rate of a new belt/component is 200 euro/hr, then after 3 hours of usage, the remaining production rate is $200 - 3 \cdot 5 = 185$ euro/hr.) This means that after t time units, the total production loss relative to a brand new belt/component are $\int_0^t axdx = at^2/2$. The relevant costs per time unit $g(\tau)$ of replacing the the belt every τ time units are therefore:

$$g(\tau) = \frac{a\tau^2/2 + C_p}{\tau} = \frac{a\tau}{2} + \frac{C_p}{\tau}. \tag{3.1}$$

It is best to replace the belt after the amount of usage that minimizes $g(\tau)$. We find that

$$\frac{dg(\tau)}{d\tau} = \frac{a}{2} - \frac{C_p}{\tau^2}, \quad \frac{d^2g(\tau)}{d\tau^2} = \frac{2C_p}{\tau^3} > 0.$$

$g(\tau)$ is convex (because its second derivative is positive for all $\tau > 0$) so the optimal amount of usage (time) after which the belt should be replaced can be found by setting $dg(\tau)/d\tau = 0$ and solving for τ. Doing this yields the optimal replacement usage/time τ^*:

$$\tau^* = \sqrt{\frac{2C_p}{a}}. \tag{3.2}$$

Example 3.1. Consider a conveyor belt that is used 24 hours per day, 7 days per week in a factory. The speed of this belt, when it is as good as new, is 5 km/hr and products on the belt are spaced 1 meter apart. The profit for each sold product is 0.5 Euro. Unfortunately, the conveyor belt speed decreases over time by 0.05 km/hr/day, and replacing the belt (C_p) costs 9600 Euro. The production rate is $5 \cdot 1000 \cdot 0.5$ Euro/hr $= 24 \cdot 0.5 \cdot 5000 = 60,000$ Euro/day, and so $a = 0.05 \cdot 1000 \cdot 0.5 \cdot 24 = 600$ Euro/day/day. Therefore, it is best to replace the belt every $\tau^* = \sqrt{2 \cdot 9600/600} = \sqrt{32} \approx 5.66$ days. $\diamond$

3.2. Stochastic maintenance models

In traditional stochastic maintenance models, we consider the question when maintenance should be performed when the only information we have is the lifetime distribution of a component. Thoughout this section, we will use the notation in Table 3.1.

Table 3.1: Notation used in traditional stochastic maintenance models

Notation		Interpretation
C_u	:	Cost of performing unplanned corrective maintenance/replacement, $C_u > 0$
C_p	:	Cost of performing planned preventive maintenance/replacement, $0 < C_p < C_u$
C_{mr}	:	Cost of performing an unplanned minimal repair. $0 < C_{mr} < C_u$
T	:	Random variable denoting the lifetime of a component.
$F_X(\cdot)$	:	Distribution function of a random variable X, $F_X(t) = \mathbb{P}(X \leq t)$
$f_X(\cdot)$	:	Density function of a random variable X, we assume $f_X = dF_X(t)/dt$
$h_X(\cdot)$	:	Failure rate of the random variable X
$M_X(\cdot)$	:	Renewal function for a renewal process with inter-renewal times that have the distribution of X.

The models in this section are usually applied to construct maintenance programs for components. In practice, maintenance programs for components will be linked to a complete maintenance program for the equipment as a whole. This integration step is not treated here, but treated later in the course. Maintenance in this case can also be replacement. We will use these terms interchangeably.

3.2.1 Failure based policy

The failure based policy is to replace/maintain a component every time it fails. This is the same as the breakdown corrective maintenance policy described in the introduction; see Table 1.1. The costs up to time t for this policy constitute a renewal reward process. The expected cycle length in this case is just the mean time to failure, $\mathbb{E}[T]$ and the expected cycle costs under this policy are

C_u. Therefore the expected costs per time unit, g, can be found by a straightforward application of the renewal reward theorem:

$$g = C_u/\mathbb{E}[T]. \tag{3.3}$$

The failure based policy is optimal when T is DFR (or CFR) because performing preventive maintenance does not actually improve the reliability of a component. Components with DFR are usually electronics. However, even if T is IFR, the failure based policy can be attractive if the failure rate does not grow without bound ($\lim_{x\to\infty} h_T(x) < \infty$) and $C_u - C_p$ is sufficiently small.

Example 3.2. Suppose that the lifetime of a components is uniformly distributed from 10 to 20 time units and that unplanned corrective maintenance costs 1000 Euro. If we choose to apply a failure based policy for this component, then the costs per time unit are $C_u/\mathbb{E}[T] = 1000/((20+10)/2) = 66.67$ Euro per time unit. (Note that the uniform distribution is IFR with $\lim_{x\to b} h_T(x) = \infty$ so it is probably a bad idea to apply failure based maintenance for this component. We consider applying preventive maintenance for this component in the next example.) ◇

3.2.2 Age replacement policy

Under the age replacement policy, a component is replaced whenever it has been used for a fixed amount of time τ or if it fails before this time. This policy fits in the upper left box of Table 1.1 from the introduction. We let the random variable $X = \min(\tau, T)$ denote the time until a component is replaced. When a components is replaced after τ amount of time, the cost of a planned replacement C_p is incurred. The unplanned maintenance cost $C_u > C_p$ is incurred if the component is replaced due to failure before τ amount of usage. We assume the time to replace is negligible. When $\tau \to \infty$, this policy is equivalent to the failure based policy. An important observation for this policy is that replacement times are also renewal points. Therefore, we can define the time until replacement as a cycle, and then the average cost under this policy can be studied as a renewal reward process. The inter-renewal time of this renewal reward process is distributed as X and the cost per cycle, W, is C_p with probability $1 - F_T(\tau)$ and C_u with probability $F_T(\tau)$. (Note that X and W are correlated random variables.) The ECC and ECL for the age replacement policy are:

$$ECC = F_T(\tau)C_u + (1 - F_T(\tau))C_p \tag{3.4}$$

$$ECL = \mathbb{E}[X] = \mathbb{E}[\min(T, \tau)] \tag{3.5}$$

$$= \int_0^\infty \min(x, \tau) f_T(x)dx = \int_0^\tau x f_T(x)dx + \tau R(\tau), \tag{3.6}$$

and the costs, which depend on τ, are $g(\tau) = ECC/ECL$. The optimal replacement time τ^* can be found by settig $dg(\tau)/d\tau = 0$ and solving for τ. If T is DFR, then $dg(\tau)/d\tau > 0$ for all τ and $\lim_{\tau\to\infty} dg(\tau)/d\tau = 0$. That is, the optimal age replacement policy reduces to a failure based policy.

Example 3.3. Reconsider the component from Example 3.2 with $C_u = 1000$, and a lifetime that is uniformly distributed between 10 and 20 time units. Suppose that planned preventive maintenance for this component costs 600 Euro. Note that under an age replacement policy only makes sense for $\tau \in (10, 20)$. For $\tau \in (10, 20)$ we have

$$ECC = F_T(\tau)C_u + (1 - F_T(\tau))C_p = \frac{\tau - 10}{10}1000 + \frac{20 - \tau}{10}600 = 40\tau + 200 \tag{3.7}$$

and

$$\begin{aligned} ECL &= \frac{1}{10}\int_{10}^{\tau} x dx + \tau\frac{20 - \tau}{10} \\ &= \frac{1}{10}\left[x^2/2\right]_{x=10}^{x=\tau} + \tau\frac{20 - \tau}{10} \\ &= \frac{1}{10}\frac{1}{2}\tau^2 - \frac{1}{10}\frac{1}{2}10^2 + 2\tau - \frac{1}{10}\tau^2 \\ &= -\frac{1}{20}\tau^2 + 2\tau - 5. \end{aligned} \tag{3.8}$$

Combining (3.7) and (3.8) yields for the expected cost per time unit (after some algebra):

$$g(\tau) = ECC/ECL = \frac{-800(\tau + 5)}{\tau^2 - 40\tau + 100}, \qquad \frac{dg(\tau)}{d\tau} = \frac{800(\tau^2 + 10\tau - 300)}{(\tau^2 - 40\tau + 100)^2}. \tag{3.9}$$

Now $dg(\tau)/d\tau = 0$ only if its numerator is 0 so the the optimal replacement time τ^* can be found by solving the quadratic equation $\tau^2 + 10\tau - 300 = 0$. Thus the optimal replacement time is after $\tau^* = 13.0278$ time units and the average cost per time unit under this policy is $g(\tau^*) = 57.37$ Euro per time unit. (In Example 3.2, we already noted that preventive maintenance should be useful in this case because the failure rate grows without bound. The saving compared to the failure based policy is $(66.67 - 57.37)/66.67 \approx 14\%$.) ◇

Example 3.4. A machine has an Erlang distributed lifetime with shape parameter $k = 2$ and scale parameter $\lambda = 1$ (per year). So, the failure distribution is given by

$$F_T(x) = 1 - (1 + x)\,e^{-x}\,, \qquad x \geq 0 \ (x \text{ in years}).$$

The costs of a preventive maintenance action are equal to 500 Euro. For a corrective maintenance action, the costs are 7000 Euro because unplanned maintenance interrupts the production process in which the machine functions. The maintenance manager would like to be able to evaluate the costs of different age based policies for this machine. Using the theory we built so far we find

$$g(\tau) = \frac{7000(1-(1+\tau)e^{-\tau}) + 500(1+\tau)e^{-\tau}}{\int_0^\tau x^2 e^{-x}dx + (\tau^2+\tau)e^{-\tau}}. \tag{3.10}$$

Working out the integral in the denominator using integration by parts twice, we have

$$\begin{aligned}\int_0^\tau x^2 e^{-x}dx &= \left[-x^2e^{-x}\right]_0^\tau + 2\int_0^\tau xe^{-x}dx \\ &= -\tau^2e^{-\tau} + 2\left([-xe^{-x}]_0^\tau + \int_0^\tau e^{-x}dx\right) \\ &= -\tau^2e^{-\tau} + 2\left(-\tau e^{-\tau} + [-e^{-x}]_0^\tau\right) \\ &= -\tau^2e^{-\tau} - 2\tau e^{-\tau} - 2e^{-\tau} + 2 \\ &= -e^{-\tau}(\tau^2+2\tau+2) + 2 \end{aligned} \tag{3.11}$$

Now substituting (3.21) back into (3.20), we find:

$$g(t) = \frac{7000(1-(1+\tau)e^{-\tau}) + 500(1+\tau)e^{-\tau}}{-e^{-\tau}(\tau^2+2\tau+2) + 2 + (\tau^2+\tau)e^{-\tau}} = \frac{7000 - 6500(1+\tau)e^{-\tau}}{-e^{-\tau}(\tau+2)+2}. \tag{3.12}$$

◇

3.2.3 Block replacement policy

Under a block replacement policy, a component is replaced at fixed times $\tau, 2\tau, 3\tau, \cdots$ and an unplanned corrective maintenance is done if the component fails between these times. This policy can be attractive when the preventive maintenance of multiple components of the same system can be coordinated. Consider for example a wind-turbine park at sea. It is beneficial to maintain all wind-turbines together at fixed times $\tau, 2\tau, 3\tau, \cdots$ because the cost of transporting goods and maintenance engineers out to sea only need to be incurred once for all wind-turbines together. The block replacement policy is a usage based policy; see the upper left of Table 1.1 in the introduction.

In the previous sections, we always defined a cycle by the time between replacements because this led to a renewal (reward) process. Under a block policy, the time at which replacements are done, are *not* renewal points. To see why, consider a component that has just been preventively replaced at time 0. The time until the next replacement is distributed as $\min(\tau, T)$. The time until the next replacement (after the first) is not distributed as $\min(\tau, T)$ unless the first replacement was a planned preventive replacement. Since the first replacement could also have been an unplanned corrective replacement, the sequence of times between replacement is not iid. Fortunately, renewal reward theory also applies to the block replacement policy by defining a cycle as the time between planned preventive replacement. Thus:

$$ECL = \tau.$$

The costs during such a cycle are C_p for the planned preventive replacement during each cycle and the costs of corrective maintenance during such a cycle. Recall that $M_T(t)$ is the expected number of renewals (failures) during an interval of length t. Therefore we have

$$ECC = C_p + C_u M_T(\tau).$$

The average costs per time unit depend on τ and are given by $g(\tau) = ECC/ECL$. Most of the time (but not always), the optimal block-replacement interval τ^* can be found by setting $dg(\tau)/d\tau = 0$ and solving for τ.

Example 3.5. Reconsider the component from Examples 3.2 and 3.3 with $C_p = 600$, $C_u = 1000$ and T has a uniform distribution from 10 to 20. Observe first that since $10 \leq T \leq 20$, $\tau^* \in (10, 20)$. (Why?) To determine the ECC, we need to find $M_T(t)$ for $t \in [10, 20]$. The number of failures in an interval of length less than 20 can be at most 1. (Why?) Therefore, $M_T(t) = F_T(t)$ for $t \in (0, 20)$; this can be verified directly using Theorem 2.1. Now it is straightforward that for $\tau \in [10, 20]$

$$ECC = C_p + C_u M_T(\tau) = 600 + 1000 F_T(\tau) = 600 + 1000\frac{\tau - 10}{10} = 100\tau - 400, \quad ECL = \tau, \tag{3.13}$$

so that the average costs per time unit are

$$g(\tau) = ECC/ECL = \frac{100\tau - 400}{\tau} = 100 - \frac{400}{\tau}, \quad \frac{dg(\tau)}{d\tau} = \frac{400}{\tau^2}. \tag{3.14}$$

Note that $g(\tau)$ is increasing because $dg(\tau)/d\tau > 0$ for $\tau \in (10, 20)$. Therefore the optimal block-replacement interval $\tau^* = 10$ time units and $g(\tau^*) = 60.00$ Euro per time unit. ◇

3.2.4 Block replacement policy with minimal repair

A block replacement policy with minimal repair is identical to a regular block replacement policy with one exception: When a component fails between block replacements, it is repaired to a state that is not as good as new, but statistically identical to the state just before the failure. Such a repair is called a minimal repair and its costs are denoted C_{mr}. In practice, minimal repairs are performed by using duct-tape, tie-wraps and other ad-hoc solutions to get a component functioning again, without actually replacing it or performing thorough maintenance. A "full" repair brings the failure rate just after replacement back to $h_T(0)$. By contrast, after minimal repair, the failure rate does not change at all. Because of this, the expected number of failures between block-replacements is no longer given by $M_T(\tau)$. Recall that the failure rate can be interpreted as the expected number of failures per time unit in a very small time interval. Therefore, $\int_0^\tau h_T(x)dx$ is the expected number of failures during a block-replacement interval. Thus we have

$$ECC = C_p + C_{mr} \int_0^\tau h_T(x)dx, \quad ECL = \tau, \tag{3.15}$$

and the expected costs per time unit are $g(\tau) = ECC/ECL$.

Example 3.6. Reconsider Examples 3.2, 3.3, and 3.5. Suppose that the costs of a minimal repair are given by $C_{mr} = 400$ Euro. As in Example 3.5, an optimal block replacement interval must be between 10 and 20 time units long. We have for $t \in (10, 20)$:

$$\begin{aligned} ECC &= C_p + C_{mr} \int_0^\tau h_T(x)dx \\ &= 600 + 400 \int_{10}^\tau \frac{1}{20-x}dx \\ &= 600 - 400 \int_{20-10}^{20-\tau} \frac{1}{u}du \\ &= 600 - 400 \left[\ln(u)\right]_{10}^{20-\tau} \\ &= 600 - 400 \left(\ln(20-\tau) - \ln(10)\right), \end{aligned}$$

where the third equality follows from using the substitution $u = 20 - x$. The costs per time unit and its first derivative for $\tau \in [10, 20)$ are given by

$$g(\tau) = \frac{600 - 400\left(\ln(20-\tau) - \ln(10)\right)}{\tau}, \quad \frac{dg(\tau)}{d\tau} = \frac{\frac{400\tau}{20-\tau} - 600 + 400(\ln(20-\tau) - \ln(10))}{\tau^2}.$$

Setting the numerator of $dg(\tau)/d\tau$ to 0 and using some algebra, we find that the optimal block replacement interval τ^* satisfies the first order condition

$$\frac{\tau^*}{20-\tau^*} + \ln(20-\tau^*) - \ln(10) - \frac{3}{2} = 0 \tag{3.16}$$

Unfortunately (3.16) cannot be solved in closed form, but a bisection search (or any other numerical root finding procedure) yields $\tau^* = 13.00$ time units so that $g(\tau^*) = 57.1$ Euro per time unit. ⋄

3.2.5 Periodic usage based replacement policy with minimal repair

For the regular age replacement policy, we assumed that we can perform preventive maintenance at any time. In practice, not all times are equally convenient. For example, in a production line, it may be preferable to only perform maintenance during the last weekend of the month, when there is no planned production. Similarly, it might be common practice to shut down a plant or production line every three months and perform all preventive maintenance during these planned down-times. (We will come back to this issue in Chapter 6.) For aircraft, it is common to visit the maintenance hangar only during certain planned down-times that are roughly equidistant. We shall refer to these planned down-times, that can be used for planned preventive maintenance, as *scheduled downs*. We assume that scheduled downs are τ units apart, i.e., there is an opportunity to perform planned preventive maintenance at times $\tau, 2\tau, 3\tau, 4\tau, \ldots$; see Figure 3.2. (Note that τ is *not* a decision variable at this point, but a known problem parameter.) We propose the following policy for the settings that we described above: After the preventive replacement of a component, we replace it preventively again at the n-th maintenance opportunity. (Note that $n \in \mathbb{N}$.) If it fails before the n-th scheduled down since the last replacement, we apply a minimal repair and replace preventively at the next scheduled down. (Note that n is a decision variable with which we can optimize the performance of a periodic age based policy with minimal repair.) A schematic representation of a sample path under such a policy is given in Figure 3.2. Planned preventive

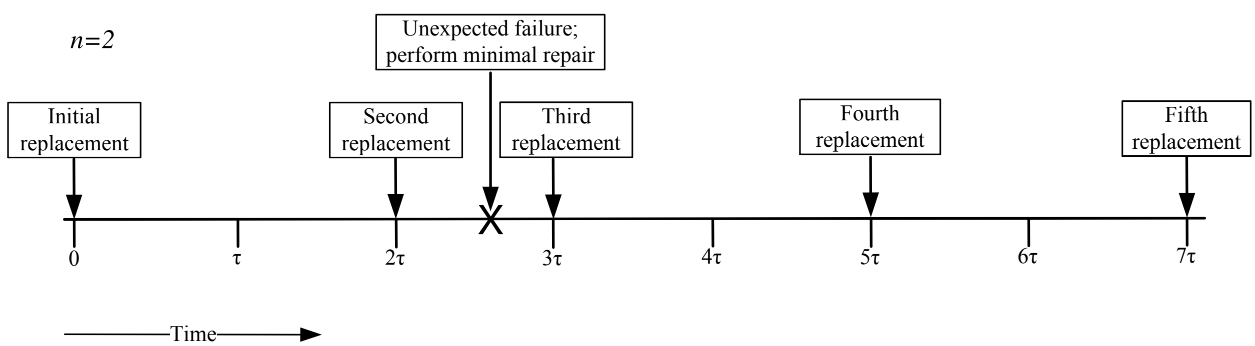


Figure 3.2: Acquisition costs versus maintenance costs over the lifetime for several assets

maintenance incurs a cost of C_p and minimal repair incur a cost of C_{mr}. If a component fails between scheduled downs, then the replacement at the next scheduled down incurs a cost C_u. The cost incurred by a periodic usage based policy with minimal repair can be described by a renewal reward process. The renewal points are the moments where a component is replaced during a scheduled down. The expected cycle length for a periodic age replacement policy is found by conditioning on the interval in which a component fails for the first time:

$$\begin{aligned} ECL &= \sum_{k=1}^{n} k\tau \mathbb{P}\left((k-1)\tau \leq T \leq k\tau\right) + n\tau \mathbb{P}(T \geq n\tau) \\ &= \sum_{k=1}^{n} k\tau \left[F_T(k\tau) - F_T((k-1)\tau)\right] + n\tau(1 - F_T(n\tau)). \end{aligned} \tag{3.17}$$

It is more involved to determine the cost during a renewal cycle. The crux is finding the expected number of minimal repairs that will be conducted during a renewal cycle. For the exposition, it is useful to introduce the random variable $Y(t_1, t_2)$ that denotes the number of minimal repairs in the interval $[t_1, t_2)$ that belong to a renewal cycle that started at time 0 ($t_1 < t_2 < n\tau$). The expected number of minimal repairs in the interval $[0, \tau)$ is given by $\mathbb{E}[Y(0, \tau)]$ under this new definition. We already know that $\mathbb{E}[Y(0, \tau)] = \int_{x=0}^{\tau} h_T(x)dx$. However, such an expression does not hold for the interval $[\tau, 2\tau)$., i.e., $\mathbb{E}[Y(\tau, 2\tau)] \neq \int_{x=\tau}^{2\tau} h_T(x)dx$. Therefore we now show a way to compute $\mathbb{E}[Y((k-1)\tau, k\tau)]$ for $k \in \mathbb{N}$. We condition on whether a failure has or has not yet occurred at the scheduled down occurring at time $(k-1)\tau$.

$$\begin{aligned} \mathbb{E}[Y((k-1)\tau, k\tau)] &= \mathbb{E}[Y((k-1)\tau, k\tau)|T > (k-1)\tau]\mathbb{P}(T > (k-1)\tau) \\ &\quad + \mathbb{E}[Y((k-1)\tau, k\tau)|T \leq (k-1)\tau]\mathbb{P}(T \leq (k-1)\tau) \\ &= \mathbb{E}[Y((k-1)\tau, k\tau)|T > (k-1)\tau](1 - F_T((k-1)\tau)) \\ &= (1 - F_T((k-1)\tau)) \int_{x=(k-1)\tau}^{k\tau} h_T(x)dx. \end{aligned} \tag{3.18}$$

The second equality above holds because $\mathbb{E}[Y((k-1)\tau, k\tau)|T \leq (k-1)\tau] = 0$ under the periodic age replacement policy. To see this, observe that if the first failure occurs before time $(k-1)\tau$, then a full replacement will occur at of before the scheduled down at time $(k-1)\tau$ so that the renewal cycle has already ended at of before time $(k-1)\tau$.

We can now determine the expected cost in a renewal cycle:

$$ECC = C_p\mathbb{P}(T \geq n\tau) + C_u\mathbb{P}(T < n\tau) + C_{mr}\sum_{k=1}^{n}\mathbb{E}[Y((k-1)\tau, k\tau)], \tag{3.19}$$

so that the expected costs per time unit are $g(n) = ECC/ECL$ with ECC given in (3.19) and ECL given in (3.17).

The most common approach to optimize the periodic usage based policy with minimal repair (minimize $g(n)$) is to compute $g(n)$ starting from $n = 1$ and continuing until you found the optimal n. Note that this approach can be streamlined computationally because for a unit increase in n, the ECC and ECL change mostly because of one additional term in the summations. $g(n)$ is not in general convex in n, but it is uni-modal in many cases of practical interest. A good approach therefore is to compute $g(n)$ until $g(n) > g(n-1)$ and declare $n^* = n - 1$.

Remark 3.1. The periodic age based policy with minimal repair reduces to the block policy with minimal repair for $n = 1$. (This holds because, as observed before, $\mathbb{E}[Y(0,\tau)] = \int_{x=0}^{\tau} h_T(x)dx$.) Furthermore, if τ is small and the n is large, then this policy will almost behave like an age based policy with age replacement threshold $n\tau$, and that model may be used.

Remark 3.2. If failed units receive full unplanned corrective maintenance, then the periodic age based policy is no longer a renewal process and exact expressions for its performance cannot be obtained. It is possible to approximate such policies using ideas similar to those in Chapter 4 of Zhu (2015).

Example 3.7. Reconsider examples 3.2, 3.3, and 3.5, i.e. $C_p = 600$, $C_u = 1000$, $C_{mr} = 400$, and T is uniformly distributed between 10 and 20. Let $\tau = 2$ (and recall that this is a parameter in the present context). The first thing to note is that $n^* \geq 5$ because the component cannot fail before its $10 = 5 \cdot 2$ time units old. Clearly for $n = 5$ we have $g(5) = 600/10 = 60$ EURO per time unit.

For $n = 6$ we have:

$$
\begin{aligned}
ECC(6) &= 1000\frac{2}{10} + 600\frac{8}{10} + 400\int_{x=10}^{12} \frac{1}{20-x}dx \\
&= 680 - 400\int_{u=20-10}^{u=20-12} \frac{1}{u}du \\
&= 680 - 400(\ln(8) - \ln(10)) = 769.26,
\end{aligned}
$$

and $ECL(6) = 12$ so that $g(6) = 769.26/12 = 64.11$ EURO per time unit.

For $n = 7$ we have:

$$
\begin{aligned}
ECC(7) &= 1000\frac{4}{10} + 600\frac{6}{10} + 400\left(\int_{x=10}^{12} \frac{1}{20-x}dx + \frac{8}{10}\int_{x=12}^{14} \frac{1}{20-x}dx\right) \\
&= 760 + 400\left(\ln(10) - \ln(8) + \frac{8}{10}(\ln(8) - \ln(6))\right) = 941.32
\end{aligned}
$$

and $ECL(7) = 12 \cdot \frac{2}{10} + 14 \cdot \frac{8}{10} = 13.6$ so that $g(7) = 69.21$ EURO per time unit. Clearly $n = 5$ is optimal in this case and the policy is equivalent to a block policy. ◇

Example 3.8. Consider a component subject to a periodic usage based replacement policy with cost parameters $C_p = 1000$, $C_u = 1500$ and $C_{mr} = 600$ and a Weibul lifetime distribution given by $F_T(t) = 1 - \exp(-(t/50)^5)$. The expected cost per time unit for this system can be evaluated exactly in closed form (try this yourself). If we plot the expected cost per time unit for $n \in \{1, 2, 3, 4, 5\}$ and $\tau \in (0, 90)$ we obtain Figure 3.3. (Verify your own expressions with Figure 3.3.) ◇

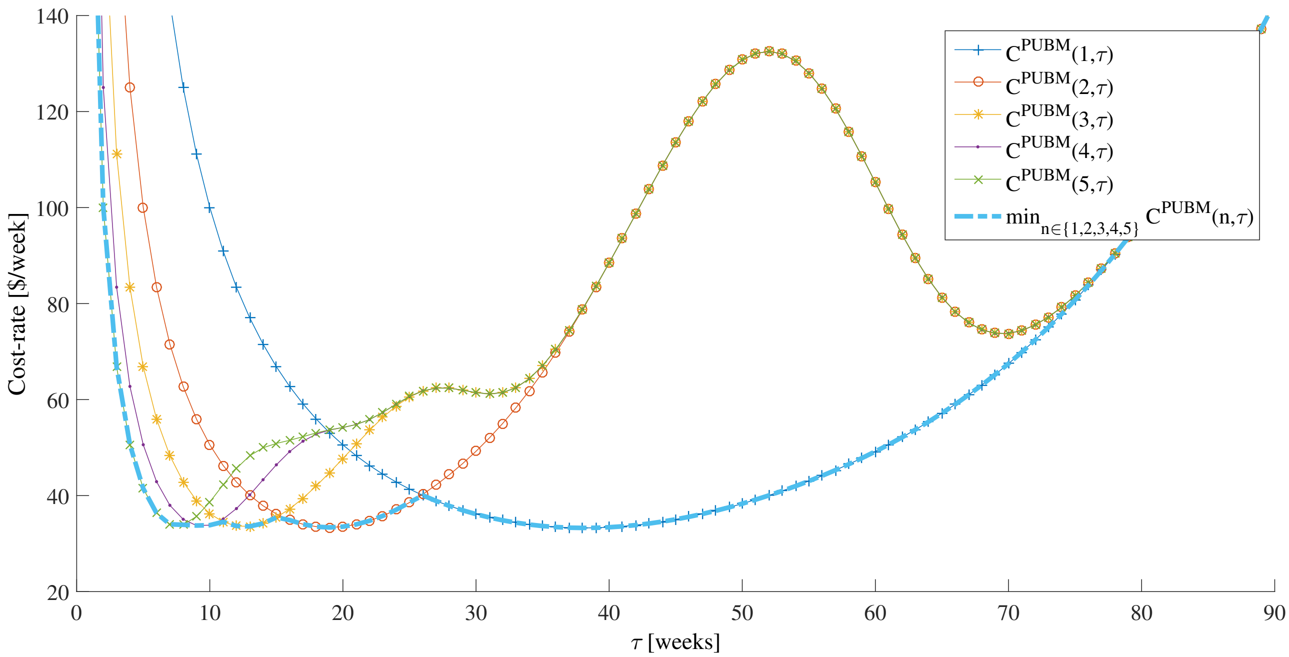


Figure 3.3: The cost of a periodic usage based maintenance policy for different values of n and optimized over $n \in \{1, 2, 3, 4, 5\}$ for a Weibul distribution with shape $\beta = 5$ and scale $\eta = 50$ with $C_p = 1000$, $C_u = 1500$, and $C_{mr} = 600$

3.3. Exercises

Exercise 3.1. *(On the age policy; level: (below) standard)*

A machine has a homogeneous (= uniform) failure distribution on the interval $[0, 10]$ (in months). So, the failure distribution is given by:

$$F(x) = \frac{x}{10}\ , \qquad 0 \leq x \leq 10\ (x \text{ in months}).$$

For this machine the following maintenance policy is used. The machine is replaced by a new one as soon as it reaches the age of τ months, or upon failure in case the machine fails before time τ. The costs of a replacement are equal to Euro 3000. For a corrective maintenance action additional costs equal to Euro 1000 are incurred because of the disturbance of the production process that depends on the availability of the machine.

Determine the average costs of the age-based maintenance policy as a function of τ and determine the optimal policy and the corresponding average costs.

Exercise 3.2. *(On the age policy; old exam exercise; level: standard)*

A machine has an Erlang distributed lifetime with shape parameter $k = 2$ and scale parameter $\lambda = 2$ (per year). So, the failure distribution is given by:

$$F(t) \; = \; 1 - (1 + 2t)e^{-2t} \; , \qquad t \geq 0 \; (t \text{ in years}).$$

(a) Determine the failure rate function. Is it useful to apply age-based preventive maintenance for this machine?

(b) The management decides that the machine has to be replaced by a new one as soon as the failure rate exceeds the value of $\alpha = 1$ per year, or as soon as the machine fails if this happens before the failure rate exceeds this level. Determine the time τ at which the machine has to be replaced preventively if the machine did not fail before.

(c) The costs of a preventive maintenance action are equal to Euro 500. For a corrective maintenance action the costs are Euro 200 higher. A maintenance engineer conjectures that the rule of the management concerning preventive replacements leads to too early replacements, and hence to higher costs than needed. He suggests that preventive replacement as soon as the failure rate exceeds the value $\alpha = 4/3$ leads to lower expected costs per year. Is he right? Is there a large difference?

Exercise 3.3. *(similar to an old exam exercise; level: above standard)*

A machine has one critical component that is subject to failures. This component has a lifetime consisting of two parts: a deterministic part with length $d \geq 0$ (in months) and a Weibull distributed part with shape parameter $\beta > 0$ and scale parameter $\eta > 0$ (in months^{-1}). The failure distribution is given by:

$$F(t) \; = \; \begin{cases} 0 & \text{if } 0 \leq t < d; \\ 1 - e^{-((t-d)/\eta)^{\beta}} & \text{if } t \geq d, \end{cases} \qquad (t \text{ in months}).$$

(a) Determine the failure rate function. In which case is the function constant, increasing, and decreasing, respectively?

(b) Is it useful to apply preventive maintenance for the component?

Assume now that $d = 0$, $\beta = 2$, and $\eta = 4$, and that the following maintenance policy is applied. If the component fails then the component itself is immediately repaired. This takes an exponential time with mean $\nu = 0.1$ months. During this time the machine is not available. After repair the component is as good as new.

(c) Determine the expected time μ till the component fails.

(d) Determine the long-run fraction of time that the machine is available.

Exercise 3.4. *(On the minimal repair policy; old exam exercise; level: above standard)*

A machine has a Weibull distributed lifetime with shape parameter $\beta = 2$ and scale parameter $\eta = 1$ (years), i.e., the failure distribution is given by:

$$F(t) \;=\; 1 - e^{-t^2} \;, \qquad t \geq 0 \;(t \text{ in years}).$$

If the machine fails then, by minimal repair, the machine can be brought back into the state it had just before the failure. The costs of such a minimal repair are equal to Euro 100. The alternative is to replace the machine by a new one, which costs Euro 900.

(a) The following maintenance policy is applied. The machine is replaced by a new one after τ_1 years. If the machine fails before the time τ_1, then minimal repair is applied. Determine the value of τ_1 which leads to the lowest average costs per year. How large are the average costs for this τ_1?

(b) Suppose now that at time τ_1 as obtained under a) the machine receives a thorough revision, which is such that the failure rate is brought back to the level of time $\tau_1 - 2$. The price of this revision is Euro 300. The time needed for this revision is small and may be neglected. Next, at time $\tau_2 > \tau_1$, the machine is replaced by a new one. Determine the value of τ_2 for which the average costs are minimal, and give the corresponding average costs.

(c) Which policy is better, the one of a) or the one of b)?

Exercise 3.5. *(On the block policy; level: above standard)*

A technical system consists of 1000 identical components. For this system, the following block maintenance policy is applied. The components are inspected at the end of each month. If a component appears to be broken, then it is immediately replaced by a new one. This costs Euro 30 per component that has to be replaced. After τ periods, $\tau \in \mathbb{N} = \{1, 2, ...\}$, all components are replaced, independently of the fact whether they are broken at that time or not. This costs Euro 10000 for the whole group of components (this is equivalent to Euro 10 per component).

Let p_i, $i \in \mathbb{N}$, denote the probability that a new components fails in the i-th month (which is observed at the end of the i-th month). These probabilities are equal to:

$$p_1 = 0,10;\ p_2 = 0,15;\ p_3 = 0,25;\ p_4 = 0,25;\ p_5 = 0,15;\ p_6 = 0,10;\ p_7 = 0 \text{ for all } i \geq 7.$$

Since the components have a maximal lifetime of 6 months, only block maintenance policies with $\tau \leq 6$ are considered.

(a) Let a cycle be defined as the time between two successive moments at which the components are replaced preventively. To determine the expected cycle costs, one has to determine the expected number of corrective maintenance actions in a cycle. For that purpose, we define M_t as the expected number of replacements after t months, $t \in \mathbb{N}$. This function may be seen as the discrete-time version of the renewal function as known from the renewal theory. For M_t, the following recursive formula holds:

$$M_t = \sum_{i=1}^{t} p_i + \sum_{i=1}^{t-1} p_i M_{t-i}\ , \qquad t \in \mathbb{N}.$$

Give an intuitive explanation for this formula.

(b) Compute M_t for $t = 1, \ldots, 6$.

(c) Give a formula for the average costs of the block maintenance policy with $\tau \in \{1, 2, \ldots, 6\}$.

(d) Determine the optimal value for τ. What are the resulting average costs.

Exercise 3.6. *(On the failure-based and minimal repair policy; old exam exercise; level: standard)* The distribution function of the lifetime of a certain machine is given by

$$F(x) = 1 - e^{-(2x)^{1.5}}\ , \qquad x \geq 0 \ (x \text{ in years}).$$

If the machine fails, then one can either replace the machine by a new one or apply a minimal repair, after which the machine is back in the same shape as just before its breakdown. The costs of a minimal repair are equal to 2000 Euro, while the costs of replacement by a new machine equal 5000 Euro.

For this machine two maintenance policies are considered: (i) The failure-based maintenance policy; (ii) The minimal repair policy, under which the machine is replaced by a new one after each τ time units and minimal repair is applied when a breakdown occurs.

(a) What are the average costs obtained under the failure-based maintenance policy? (Given: $\Gamma(1.667) = 0.9033$.)

(b) Determine the failure rate function. What can be concluded from the behavior of the failure rate function for the minimal repair policy?

(c) Determine the minimal average costs that can be obtained by the minimal repair policy. How large is the difference with respect to the average costs obtained for the failure-based maintenance policy? Are you surprised by this difference?

Exercise 3.7. *(On the age maintenance policy; old exam exercise; level: above standard)*

A dealer of an automatic orange juice press currently has about 100 customers. These customers are mainly restaurants and bars and each of them has one press.

When a customer buys a new press, he receives a (free) warranty of the manufacturer for failures within the first 3 years. If a failure occurs within the first 3 years, then the maintenance department of the manufacturer will replace the failed press by a new one for free. It is known that the presses are very reliable in the first 3 years, and hence these free replacements occur seldom.

After 3 years, a press can break down at any time. Statistics from the past suggest that the remaining lifetime after 3 years is exponential with mean 1.5 years. As a result, the lifetime distribution of a new press may be approximated by the following distribution:

$$F(t) \;=\; \begin{cases} 0 & \text{if } 0 \le t < 3; \\ 1 - e^{-\frac{2}{3}(t-3)} & \text{if } t \ge 3, \end{cases} \qquad (t \text{ in years}).$$

If the press fails after the first 3 years, then a customer will buy a new press at the dealer, for which he will receive a warranty of 3 years again. The price of a new press is Euro 500 if it is ordered by the normal procedure. This procedure is used when the press is replaced by a new one preventively. If a press is only replaced after it has failed, then a new press is ordered and delivered by the emergency procedure, which leads to additional costs equal to C_{emer} (in Euro). These costs are subject to a change and hence C_{emer} is seen as a variable now.

(a) Determine the failure rate function for the orange juice press. Can you conclude from this function whether it is sensible, or not, to apply preventive replacements?

Suppose that at each customer, replacements of the press are applied according to an age policy with parameter τ. The corresponding average costs per year are denoted by $g(\tau)$. We also assume that the customers are rational, i.e. that the τ is chosen such that the average costs $g(\tau)$ are minimized.

(b) Determine the average costs $g(\tau)$.

(c) Determine the optimal age policy.

(d) Currently a customer has to pay $C_{emer} = 200$ Euro extra when he orders a new press by the emergency procedure instead of the standard procedure. The dealer plans to charge 300 Euro extra from now on. What are the consequences of this increase for both the customers and the dealer?

Exercise 3.8. *(On the age policy; old exam exercise; level: (above) standard)*

A machine has an Erlang distributed lifetime with shape parameter $k = 2$ and scale parameter $\lambda = 1$ (per year). So, the failure distribution is given by:

$$F(t) \; = \; 1 - (1+t)e^{-t} \;, \qquad t \geq 0 \; (t \text{ in years}).$$

(a) Determine the mean and standard deviation of the lifetime of the machine.

(b) Determine the failure rate function. Is it useful to apply preventive maintenance for this machine?

The costs for a preventive maintenance action for this machine are equal to 200 Euro. The costs for a corrective maintenance action are equal to $200 + C_{emer}$ Euro, where the factor $C_{emer} \geq 0$ represents emergency costs. The higher C_{emer} the more attractive it is to apply preventive maintenance. We now consider an age maintenance policy with parameter τ.

(c) Determine the average costs $g(\tau)$ for the age policy with parameter τ.

(d) The derivative of $g(\tau)$ is denoted by $g'(\tau)$. Show that: $g'(\tau) = 0$ if and only if

$$[(C_{emer} - 200)\tau - (C_{emer} + 200)] + C_{emer}e^{-\tau} \; = \; 0 \; .$$

When is $g'(\tau) > 0$? When is $g'(\tau) < 0$?

(e) Intuitively, one would think that it becomes attractive to apply a preventive maintenance action when the lifetime has become equal to the mean lifetime. What is your opinion on this intuition?

Exercise 3.9. *(On the failure-based and minimal repair policy; level: standard)*

The distribution function of the lifetime of a certain machine is given by

$$F(x) \; = \; 1 - e^{-(3x)^2} \;, \qquad x \geq 0 \; (x \text{ in years}).$$

If the machine fails, then one can either replace the machine by a new one or apply a minimal repair, after which the machine is back in the same shape as just before its breakdown. The costs of a minimal repair are equal to 400 Euro, while the costs of replacement by a new machine equal 2000 Euro.

For this machine two maintenance policies are considered: (i) The failure-based maintenance policy; (ii) The minimal repair policy, under which the machine is replaced by a new one after each τ time units and minimal repair is applied when a breakdown occurs.

(a) What are the average costs obtained under the failure-based maintenance policy? (Given: $\Gamma(1.5) = 0.8862$.)

(b) Determine the failure rate function. What can be concluded from the behavior of the failure rate function for the minimal repair policy?

(c) Determine the minimal average costs that can be obtained by the minimal repair policy. How large is the difference with respect to the average costs obtained for the failure-based maintenance policy? Are you surprised by this difference?

Exercise 3.10. *(On the failure-based and age policy; old exam exercise; level: above standard)*
In this exercise we consider a component with a uniform (= homogeneous) failure distribution. We assume that the part may fail from the beginning. Without loss of generality, we assume that the maximum lifetime is 1. Hence the failure distribution $F(t)$ is given by $F(t) = t$, $0 \leq t \leq 1$. The component is part of an important machine. As soon as the component fails it has to be replaced by a new one.

For this component we investigate two maintenance policies, the failure-based policy and the age policy with parameter τ, $0 \leq \tau \leq 1$. The costs for a planned replacement of the component are denoted by C_p. The costs for an unplanned replacement are equal to C_u. We assume that $C_u > C_p > 0$.

The average costs under the failure-based policy are denoted by g_{fb}. The average costs under the age policy with parameter τ are denoted by $g_{age}(\tau)$. The parameter for which these costs are minimized is denoted by τ^* and $g^*_{age} = g_{age}(\tau^*)$. The ratio $(g_{fb} - g^*_{age})/g_{fb}$ denotes the relative savings that are obtained under the optimal age policy in comparison to the failure-based policy. We are interested in how large these savings are as a function of the ratio C_u/C_p.

(a) Determine g_{fb}.

(b) Determine $g_{age}(\tau)$, $0 < \tau \leq 1$.

(c) Determine the derivative $g'_{age}(\tau)$ of $g_{age}(\tau)$ and show that

$$g'_{age}(\tau) = 0 \quad \Leftrightarrow \quad \frac{1}{2}\left(\frac{C_u}{C_p} - 1\right)\tau^2 + \tau - 1 = 0 \ ,$$

and similarly with the "="-sign replaced by the ">"- and "<"-sign, respectively.

(d) Show that

$$\tau^* = \frac{\sqrt{1 + 2\left(\frac{C_u}{C_p} - 1\right)} - 1}{\left(\frac{C_u}{C_p} - 1\right)}.$$

(e) Compute $(g_{fb} - g^*_{age})/g_{fb}$ for $\frac{C_u}{C_p} \downarrow 1$ and for $\frac{C_u}{C_p} = 2,\ 5,\ 13$.

(f) Give your opinion on the relative savings that are possible by the age policy in comparison to the failure-based policy.

Exercise 3.11. *(On block replacement, above standard)*

A wind-turbine park in the north-sea consists of 10 identical turbines. Each month, a maintenance engineer visits the park and repairs turbines that have failed. Such a repair brings them to as good as new condition and costs of 500 € per turbine. Once every τ months, a team of engineers visits the wind-turbine park to maintain *all* turbines and return them to as good as new condition. This preventive maintenance costs 2000 € for the entire park. The lifetime of any turbine has a right-truncated Weibull distribution with $\beta = 2$ and $\eta = 5$ and truncation parameter $R = 12$, i.e., the probability a turbine fails before x months of service is given by:

$$F(x) = \begin{cases} 1 - \exp\left(-(x/\eta)^\beta\right), & \text{if } x < R; \\ 1, & \text{if } x \geq R. \end{cases}$$

This means that the lifetime of a turbine has a Weibull distribution up to time R, but is certain to fail at time R if it has not failed before time R. Let $p_i = F(i) - F(i-1)$ for $i \in \{1, \ldots, 12\}$

(a) Suppose that preventive maintenance is never conducted ($\tau = \infty$). How many visits will the maintenance engineer pay to an arbitrary turbine until he finds that it has failed? Give an expression in terms of p_i and compute the numerical answer.

(b) Now suppose that we want to determine the monthly cost as a function of τ. To determine the expected cycle costs, one has to determine the expected number of corrective maintenance actions in a cycle. For that purpose, we define M_t as the expected number of turbine repairs after t months, $t \in \mathbb{N}$. This function may be seen as the discrete-time version of the renewal function as known from the renewal theory. For M_t, the following recursive formula holds:

$$M_t = \sum_{i=1}^{t} p_i + \sum_{i=1}^{t-1} p_i M_{t-i}\ , \qquad t \in \mathbb{N}.$$

Give an intuitive explanation for this formula.

(c) Compute M_t for $t = 1, \ldots, 12$.

(d) Give a formula for the average monthly cost of the block maintenance policy with $\tau \in \{1, 2, \ldots, 12\}$.

(e) Determine the optimal value for τ. What is the resulting average monthly cost?

Exercise 3.12. *(Age replacement and failure rate, standard)*

A machine has an Erlang distributed lifetime with shape parameter $k = 2$ and scale parameter $\lambda = 1$ (per year). So, the failure distribution is given by

$$F(x) = 1 - (1 + x)\, e^{-x} \,, \qquad x \geq 0 \text{ (x in years).}$$

We assume that the machine is needed for an infinite horizon. Corrective maintenance is executed when the machine fails, after which the machine is as good as new. Preventive maintenance is also possible. Also after preventive maintenance, the machine is as good as new. The costs of a preventive maintenance action are equal to 500 €. For a corrective maintenance action, the costs are 7000 € because unplanned maintenance interrupts the production process in which the machine functions.

(a) Currently a failure-based policy is followed. Determine the yearly average costs under this policy.

(b) Could it be useful to apply preventive maintenance for this machine? Why (not)?

(c) We consider an age policy as an alternative for the failure-based policy. Assume that preventive maintenance is executed as soon as the failure rate (also called hazard rate) reaches the value of $\frac{1}{6}$ failures per year. If the machine fails before that time point, then corrective maintenance is executed. Determine the parameter τ_0 of the resulting age policy (i.e., the policy that executes preventive maintenance when the component age reaches τ_0 years).

(d) Let $g(\tau)$ be the average yearly costs of the age policy with parameter τ. Derive the formula for $g(\tau)$.

(e) Determine the average yearly costs $g(\tau_0)$ of the age policy as described under c). How large is the difference with respect to the average cost obtained for the failure-based policy? Do you find this difference surprising? Why (not)?

Exercise 3.13. *(All policies, Old exam, above standard)*

We consider 12 solar-energy collectors (collector for short) that are stationed in the desert. The lifetime of each collector is independently and uniformly distributed between 4 and 8 months.

To perform maintenance on these collectors, an engineer needs to drive into the desert. A planned visit costs 2750 Euro and an unplanned visit costs 5000 Euro. There are material maintenance cost of 250 Euro per collector that is maintained preventively and 1000 Euro per collector that is maintained correctively. (You may assume that material needed for maintenance is always available). Failures of collectors are observed immediately because the voltage generated by the solar park drops after a failure occurs.

(a) Can it be beneficial to perform preventive maintenance on collectors? Why (not)?

(b) Suppose that a failure-based maintenance policy is used to maintain these collectors. What would be the annual total maintenance cost in this case?

(c) Suppose that an age based is used to maintain these collectors separately. What is the annual cost of an optimal age-based maintenance policy and when should collectors be replaced under such a policy?

(d) Management decides that it is silly to drive all the way into the desert (and pay either 3000 or 5000 Euro) to maintain only one collector. If an engineer can maintain multiple collectors during a visit, this might save money. Explain what a block policy is and how it can be used to avoid driving into the desert to fix only one collector.

(e) An engineer claims that under a block policy, the optimal time between block replacements, denoted by τ^*, satisfies $4 \leq \tau^* \leq 8$ (time measured in months). He further claims that any collector can fail at most once between block replacements. Is the engineer correct? Why (not)?

(f) Determine the annual cost and maintenance interval for an optimal block maintenance policy. (Hint: use the results from (e) to determine the renewal function.)

Exercise 3.14. *(Deterministic maintenance, above standard)*

Consider the air filter of a car. The price for replacing an air filter is 80 Euro. Suppose this car is driven 2000 km/month and petrol costs 1.50 Euro/l. With a new air filter, the car can drive 15 km/l, but this decreases by 1 km/l/month.

(a) Suppose you start with a new air filter. Show that up to time $0 \leq t \leq 15$ (in months) the total amount of money (in Euros) spent on petrol is

$$\int_0^t \frac{3000}{15 - x} dx$$

when we do not replace the filter. (Hint: What are the fuel consumption rate and the money consumption rate at any time?)

(b) Use substitution to show that

$$\int_0^t \frac{3000}{15-x}dx = 3000\ln\left(\frac{15}{15-t}\right)$$

(c) Consider the usage based policy in which the filter is replaced every τ time units. Use (a) and (b) to give an expression for the costs per time unit for a given τ, and evaluate this expression for $\tau = 1, 2, 3, 4, 5$.

(d) Suppose that you would like to replace your air filter after an integer number of months, when should you replace your air filter?

3.3.1 Solutions

We provide answers/solutions for all exercises and complete solutions for the last few exercises.

Solution 3.1. Average costs: $g(\tau) = \frac{60000+2000\tau}{20\tau-\tau^2}$ Euro per month; Optimal age-based policy: τ^* $= 8.73$ months; $g(\tau^*) = 787$ Euro per month.

Solution 3.2. (a) $h(t) = \frac{4t}{1+2t}$, $t \geq 0$.

(b) $\tau = 0.5$ years.

(c) $\alpha = 4/3$ is equivalent with $\tau = 1$ year.
$g(\tau) = \frac{700-200(1+2\tau)e^{-2\tau}}{1-(1+\tau)e^{-2\tau}}$, $\tau \geq 0$.
The engineer is right since $g(0.5) = 1234$ Euro per year < 848 Euro per year $= g(1)$; $g(1) - g(0.5) = 386$ Euro per year.

Solution 3.3. (a)

$$h(t) = \begin{cases} 0 & \text{if } 0 \leq t < d; \\ \beta\frac{(t-d)^{\beta-1}}{\eta^\beta} & \text{if } t \geq d. \end{cases}$$

$h(t) = 0$ on $[0, d)$; on $[d, \infty)$, $h(t)$ is constant if $\beta = 1$ (equal to η^{-1}), $h(t)$ is increasing if $\beta > 1$, and $h(t)$ is decreasing if $0 < \beta < 1$.

(b) Only if $d = 0$ and $0 < \beta \leq 1$, it can be said beforehand that it is not useful to apply preventive maintenance. In all other cases, i.e. if $d > 0$ or $\beta > 1$, it depends on the costs for the different maintenance actions whether preventive maintenance is useful or not.

(c) $\mu = 3.54$ months.

(d) 97.3 %

Solution 3.4. (a) $g(\tau_1) = \frac{900}{\tau_1} + 100\tau_1$ Euro per year, $\tau_1 \geq 0$.
$\tau_1^* = 3$ year; $g(\tau_1^*) = 600$ Euro per year.
(For comparison: the no preventive maintenance policy would lead to average costs equal to 1016 Euro per year.)

(b) Average costs as a function of τ_2: $\frac{2400}{\tau_2} + 100\tau_2 - 400$, $\tau_2 \geq \tau_1$.
$\tau_2^* = 2\sqrt{6} = 4.9$ years; the corresponding average costs are equal to 580 Euro per year.

(c) The policy of b) is slightly better; it is 20 Euro per year cheaper than the policy of a).

Solution 3.5. (b) $M_1 = 0.10$; $M_2 = 0.26$; $M_3 = 0.541$; $M_4 = 0.868$; $M_5 = 1.158$; $M_6 = 1.461$.

(c) $g(\tau) = \frac{10000(1+3M_{\tau-1})}{\tau}$ Euro per month. (Note that at time τ everything is replaced.)

(d) $\tau^* = 3$ months; $g(\tau^*) = 5933$ Euro per month.

Solution 3.6. (a) 11071 Euro per year.

(b) $h(x) = 3\sqrt{2} \cdot \sqrt{x}$, $x \geq 0$. The failure rate function is increasing and goes to infinity if $x \to \infty$. Hence, we will find a finite value for the optimal τ for the minimal repair policy.

(c) $\tau^* = 1.46$ years; $g(\tau^*) = 10260$ Euro per year; difference $= 811$ Euro per year.

Solution 3.7. (a)

$$h(t) = \begin{cases} 0 & \text{if } 0 \leq t < 3; \\ \frac{2}{3} & \text{if } t \geq 3; \end{cases}$$

This function is increasing, and hence preventive maintenance may be sensible.

(b) Since the failure rate is 0 on $[0, 3)$, only values for τ with $\tau \geq 3$ are interesting;

$$g(\tau) = \frac{500 + C_{emer}(1 - e^{-\frac{2}{3}(\tau-3)})}{3 + \frac{3}{2}(1 - e^{-\frac{2}{3}(\tau-3)})}, \qquad \tau \geq 3.$$

(c) By analyzing $g'(\tau)$, it may be shown that $g(\tau)$ is strictly increasing on $[3, \infty)$ if $C_{emer} > 250$ Euro, constant if $C_{emer} = 250$ Euro, and strictly decreasing if $C_{emer} < 250$ Euro. Hence, $\tau^* = 3$ if $C_{emer} > 250$ Euro, $\tau^* = \infty$ if $C_{emer} < 250$ Euro, and τ^* may be taken equal to any value in $[3, \infty)$ if $C_{emer} = 250$ Euro.

(d) In the current situation, $C_{emer} = 200$ Euro, and hence: (i) $\tau^* = \infty$ and $g(\tau^*) = 156$ Euro per year; (ii) Customers replace their orange juice press only correctively, on average once per 4.5 years; (iii) The dealer sells on average $\frac{100}{4.5} = 22.2$ presses per year.

If C_{emer} is increased to 300 Euro, then: (i) τ^* becomes equal to 3 years and $g(\tau^*) = 167$ Euro per year; (ii) Customers start to replace their orange juice preventively each 3 years; (iii) The dealer sells on average $\frac{100}{3} = 33.3$ presses per year.

Solution 3.8. (a) 2 years and 1.41 years, respectively.

(b) $h(t) = \frac{t}{t+1}$, $t \geq 0$. The failure rate function is increasing, and thus preventive maintenance may be sensible.

(c)

$$g(\tau) = \frac{200 + C_{emer}[1 - (1+\tau)e^{-\tau}]}{2 - (2+\tau)e^{-\tau}}, \qquad \tau \geq 0.$$

(d) $g'(\tau) > 0$ if and only if $[(C_{emer} - 200)\tau - (C_{emer} + 200)] + C_{emer}e^{-\tau} > 0$;
$g'(\tau) < 0$ if and only if $[(C_{emer} - 200)\tau - (C_{emer} + 200)] + C_{emer}e^{-\tau} < 0$.

(e) The following appears to hold:

- If $C_{emer} = 528$ Euro, then it is optimal to apply preventive maintenance when the lifetime has reached the mean lifetime;
- If $C_{emer} > 528$ Euro, then it is sensible to apply preventive maintenance when the lifetime has reached the mean lifetime, but earlier is even better;
- If $C_{emer} < 528$ Euro, then it is better to wait longer with applying preventive maintenance, and no preventive maintenance should be applied at all in case $C_{emer} \leq 200$.

Solution 3.9. (a) 6770 Euro per year.

(b) $h(x) = 18x$, $x \geq 0$. The failure rate function is increasing and goes to infinity if $x \to \infty$. Hence, we will find a finite value for the optimal τ for the minimal repair policy.

(c) $\tau^* = 0.75$ years; $g(\tau^*) = 5367$ Euro per year; difference = 1403 Euro per year.

Solution 3.10. (a) $g_{fb} = 2C_u$.

(b)

$$g_{age}(\tau) = \frac{C_p + (C_u - C_p)\tau}{-\frac{1}{2}\tau^2 + \tau}, \qquad 0 < \tau \leq 1.$$

(c) 0 % if $\frac{C_u}{C_p} \downarrow 1$; 6.7 % if $\frac{C_u}{C_p} = 2$; 20 % if $\frac{C_u}{C_p} = 5$; 31 % if $\frac{C_u}{C_p} = 13$.

Solution 3.11. (a) Let N be the number of visits a maintenance engineer pays to an arbitrary turbine until he finds that it has failed. Then $\mathbb{P}\{N = i\} = p_i$ and so $\mathbb{E}[N] = \sum_{i=1}^{12} ip_i = 4.9263$.

(b) The first term of M_t is the expected number of times the first turbine will fail in $(0, t)$. The second term described the expected number of turbine failures beyond the first failure. To find the expected number of turbine failures beyond the first, we condition on the failure time of the first turbine: If the first turbine fails in period i (with probability p_i), then the expected number of remaining failures is M_{t-i}.

(c) Filling in the given equation, we find:

t	1	2	3	4	5	6	7	8	9	10	11	12
M_t	0.0392	0.1494	0.3124	0.5072	0.7157	0.9262	1.1338	1.3379	1.5401	1.7419	1.9440	2.1498

(d) Note that there are 10 turbines and that turbines that fail between $\tau - 1$ and τ are maintained at a cost of 200 Euro per turbine. Let $C_u = 500$ and $C_p = 2000$ and let $g(\tau)$ be the average monthly cost for the entire park under a block replacement policy with interval τ; then we find

$$g(\tau) = \frac{10C_u M_{\tau-1} + C_p}{\tau} = \frac{5000M_{\tau-1} + 2000}{\tau}$$

(e) First observe that because a turbine always fails within 12 months the optimal $\tau \in \{1, ..., 12\}$. Thus, using the answers of (c) and (d) we find:

τ	1	2	3	4	5	6	7	8	9	10	11	12
$g(\tau)$	2000.00	1098.03	915.66	890.55	907.25	929.76	947.29	958.60	965.50	970.05	973.59	976.68

From this we conclude that the optimal τ is 4 months and that the optimal average monthly cost is 890.55 Euro.

Solution 3.12. (a) Let X be the lifetime of a machine. For a failure based policy $ECC = 7000$ Euro and since X has an Erlang distribution with $k = 2$ and $\lambda = 1$, $ECL = k/\lambda = 2$ years. Combining we find that the average yearly cost under a failure based policy is $ECC/ECL = 7000/2 = 3500$ Euro per year.

(b) It is only useful to apply preventive maintenance for machines that have an increasing failure rate. An Erlang distribution with $k \geq 2$ has an increasing failure rate and so it might be useful to apply preventive maintenance in this case.

(c) The failure rate of X is given by $h(x) = f(x)/(1 - F(x)) = x/(x + 1)$. Setting $h(x) = 1/6$ and solving for x yields $\tau_0 = 1/5 = 0.2$.

(d) Setting $C_u = 7000$ and $C_p = 500$ for an age based policy, we obtain

$$g(\tau) = \frac{7000(1 - (1 + \tau)e^{-\tau}) + 500(1 + \tau)e^{-\tau}}{\int_0^\tau x^2 e^{-x} dx + (\tau^2 + \tau)e^{-\tau}} \tag{3.20}$$

Working out the integral in the denominator using integration by parts twice, we have

$$
\begin{aligned}
\int_0^\tau x^2 e^{-x} dx &= \left[-x^2 e^{-x}\right]_0^\tau + 2\int_0^\tau x e^{-x} dx \\
&= -\tau^2 e^{-\tau} + 2\left([-xe^{-x}]_0^\tau + \int_0^\tau e^{-x} dx \right) \\
&= -\tau^2 e^{-\tau} + 2\left(-\tau e^{-\tau} + [-e^{-x}]_0^\tau \right) \\
&= -\tau^2 e^{-\tau} - 2\tau e^{-\tau} - 2e^{-\tau} + 2 \\
&= -e^{-\tau}(\tau^2 + 2\tau + 2) + 2 \qquad (3.21)
\end{aligned}
$$

Now substituting (3.21) back into (3.20), we find:

$$
g(\tau) = \frac{7000(1-(1+\tau)e^{-\tau}) + 500(1+\tau)e^{-\tau}}{-e^{-\tau}(\tau^2+2\tau+2) + 2 + (\tau^2+\tau)e^{-\tau}} = \frac{7000 - 6500(1+\tau)e^{-\tau}}{-e^{-\tau}(\tau+2)+2} \qquad (3.22)
$$

(e) Filling in (3.22), we find $g(\tau_0) = g(0.2) = 3088.15$ Euro per year. This is not surprising because C_p is much smaller than C_u and X has IFR. It is perhaps surprising that the benefit is 11.86% already, even though τ_0 is not the optimal τ for an age based policy.

Solution 3.13. (a) Yes, because the uniform distribution has increasing failure rate and because preventive maintenance is cheaper than corrective maintenance.

(b) $ECC = C_u = 5000 + 1000 = 6000$, $ECL = \mu = (8+4)/2 = 6$, Cost per collector are 6000/6=1000 Euro per month. For the whole collector park, it is $12 \cdot 1000 = 12000$ Euro per month.

(c) Let τ denote the replacement threshold. Define

$$
f(x) = \begin{cases} 1/4, & \text{if } 4 \le x \le 8; \\ 0, & \text{otherwise;} \end{cases} \qquad F(x) = \begin{cases} 0, & \text{if } x < 4; \\ (x-4)/4, & \text{if } 4 \le x \le 8; \\ 1, & \text{if } x \ge 8. \end{cases}
$$

Then for $4 \le \tau \le 8$:

$$
ECC = 6000F(\tau) + 3000(1-F(\tau)) = 6000\frac{\tau-4}{4} + 3000\frac{8-\tau}{4} = 750\tau, \qquad (3.23)
$$

$$
ECL = \int_4^\tau t\frac{1}{4}dt + \frac{8-\tau}{4}\tau = 2\tau - \tau^2/8 - 2, \qquad (3.24)
$$

$$
\qquad (3.25)
$$

Setting $g'(\tau) = \frac{d}{d\tau}\frac{ECC}{ECL} = 0$ yields the following quadratic equation for τ^*:

$$\frac{750}{8}\tau^2 - 1500 = 0$$

which yields $\tau^* = \sqrt{16} = 4.0$. This means the complete costs for the solar park are $12g(4) =$ 9000.00 Euro per month.

(d) In a block policy, all 12 solar panels are maintained preventively simultaneously every τ months. Therefore, we only need to pay the 5000 Euro to drive into the desert once for 12 collectors for each block maintenance moment.

(e) Before 4 months, collectors cannot fail and after 8 months they will all fail for sure so $4 \leq \tau^* \leq 8$. Since the time between failures is at least 4, there can happen only one failure in any interval of length less than 8.

(f) Observe (by arguments of (e)) that $M(t) = F(t)$ for $4 \leq t \leq 8$. Now we define a cycle for all 12 collectors together. We have:

$$ECC = 12M(\tau)6000 + 2750 + 12 \cdot 250 = 18000\tau - 66250, \tag{3.26}$$

$$ECL = \tau \tag{3.27}$$

$$g(\tau) = ECC/ECL = 18000 - 66250/\tau, \tag{3.28}$$

Observe that g is increasing in τ for $\tau > 0$, so $\tau^* = 4$. Therefore, the costs for maintaining the whole solar park will be 18000-66250/4=1437.50 Euro per month.

Solution 3.14. (a) At time t, the car can drive $15 - t$ km/l, so the fuel consumption rate at time t is $2000/(15-t)$ l/month. Since a liter of petrol costs 1.5 Euro/l, the money consumption rate at time t is $1.5 \cdot 2000/(15-t) = 3000/(15-t)$. The total costs incurred up to time t is the integral of the money consumption rate from 0 to t.

(b) We use the substitution $u = 15 - x$ so $-du = dx$.

$$\begin{aligned}
\int_0^t \frac{3000}{15-x}dx &= -3000\int_{15}^{15-t}\frac{1}{u}du \\
&= -3000\left[\ln(u)\right]_{u=15}^{u=15-t} \\
&= 3000(\ln(15) - \ln(15-t)) \\
&= 3000\ln\left(\frac{15}{15-t}\right)
\end{aligned}$$

(c) Let $g(\tau)$ denote the costs per time unit.

$$g(\tau) = \frac{3000}{\tau} \ln\left(\frac{15}{15-\tau}\right) + \frac{80}{\tau}$$

Now it is straightforward to compute:

τ	1	2	3	4	5
$g(\tau)$	286.9786	254.6513	249.8102	252.6162	259.2791

(d) The filter should be replaced every three months.

Chapter 4

Degradation Models

> "Essentially all models are wrong, but some are useful."
>
> George Box

We need to understand how assets and components degrade before we can use this understanding to make better maintenance decisions based on the condition, rather than the age of an asset or component. An indicator of the condition/degradation level of a component depends heavily on the technology. The condition of a break-pad, for example, is its thickness. The thickness of break-pads can be measured periodically when a vehicle enters the maintenance shop, but with modern sensing technology, it is also possible to monitor it continuously over time. In either case, we need to model the way the thickness of the break-pad evolves over time, and use data to fit such models. Modern high-tech equipment generates many signals from sensors in the equipment that monitor such things as vibration amplitude, temperature, light intensity, concentration of contaminants in lubrication fluids, deformation, and position of parts relative to each other. Figure 4.1 shows several modern sensors that are used to measure several physical quantities over time. These measurements provide information about the condition of equipment and are generated in increasing abundance. Using this data is an important field within the big-data movement, partially because this application is devoid of most privacy issues. This chapter is devoted to the development of degradation models, with a special emphasis on Markov degradation models. In Chapter 5, we will use these models to evaluate and optimize condition based maintenance policies.

There are 3 main types of degradation processes that are most commonly used in practice and literature. We treat representative models from all three classes. However, there are some

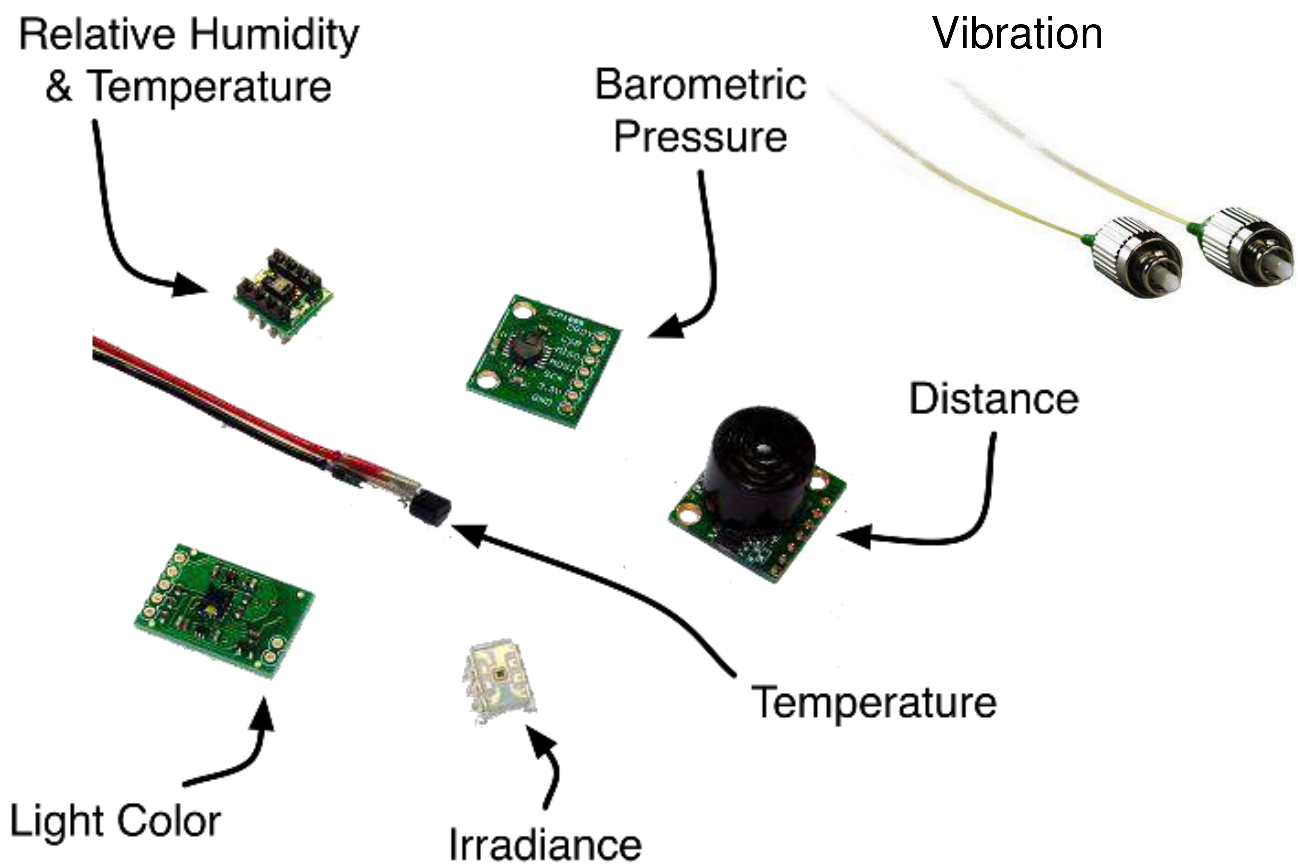


Figure 4.1: Different types of sensors that are used to monitor equipment.

common notions and notations that we introduce here. A degradation process is a stochastic process $X(t)$. We will assume for convenience that $X(0) = 0$[1]. There is a threshold L such that a component fails at time $T = \inf\{t|X(t) \geq L\}$; see Figure 4.3. We assume that $\lim_{t\to\infty} X(t) \geq L$ with probability one, i.e., a component will fail eventually. We call a stochastic process $X(t)$ that satisfies this condition a degradation process. Figure 4.3 depicts a degradation path of a roller bearing. For roller bearings (see Figure 4.2), a good condition measure is the amplitude of vibrations as measured by an accelerometer.

4.1. Random Coefficient Models

Suppose that $X(t)$ is given by some function $X(t) = f(t|\theta)$ where $\theta = (\theta_1, \ldots, \theta_n)$ are the parameters (or coefficients) of the function $f(t)$. Examples of this are $X(t) = \theta t$ and $X(t) = \theta_1 t + \theta_2\sqrt{t}$.

[1] If $X(0) = a \neq 0$, we can study the process $\tilde{X}(t) = X(t) - a$.

Figure 4.2: An example of a roller bearing.

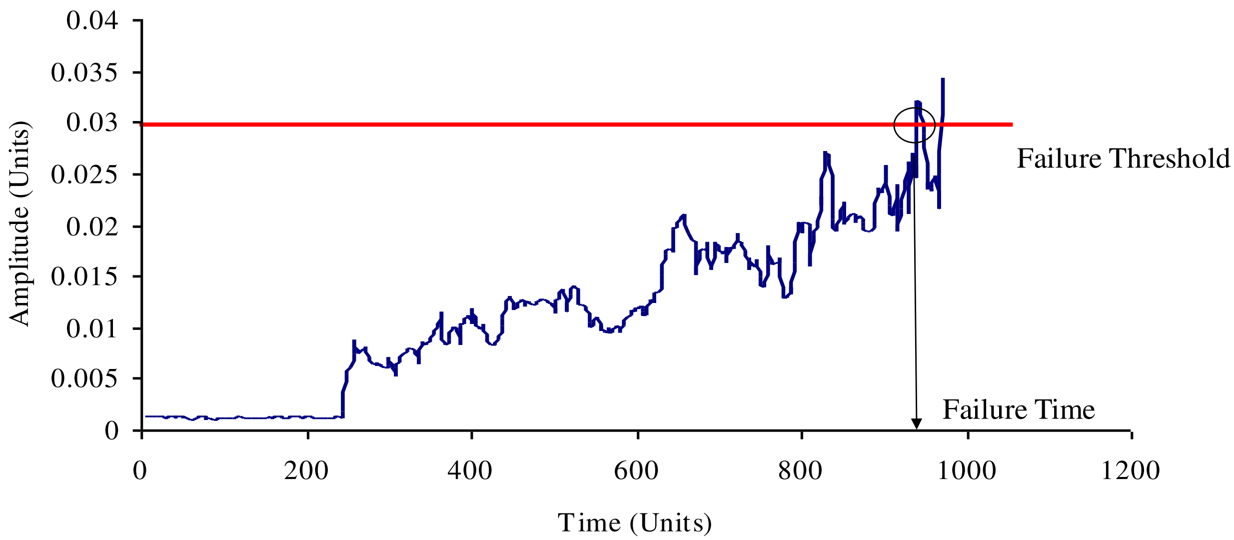


Figure 4.3: A sample degradation path of a ball bearing.

The degradation path is deterministic if θ is known (and one such case is covered in Section 3.1). However, we will assume that θ is a random variable. $X(t)$ then becomes a stochastic process, although a finite number of samples of $X(t)$ will allow you to compute the realization of θ and therefore reveal the entire future degradation path. In general, if $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$, then n samples of $X(t)$ at different points in time will give you n equations in the n unknown variables so that the future degradation path is known deterministically from then on. This is a rather strange model feature that makes random coefficient models inappropriate in maintenance models where $X(t)$ is inspected more than n times; see also van Noortwijk (2009). Despite this, the random coefficient model is used in much literature, which is why we cover its basics here. We will *not* use this degradation model to feed into CBM policies later. The original application of random coefficient models is to estimate the time-to-failure distribution based on degradation data; see Lu and Meeker (1993). We provide an example of this below.

Example 4.1. Bernstein distribution. The condition of break-pads is measured by their thickness in some appropriate unit such as mm. Suppose that $X(t)$ denotes the thickness that has been lost since a break-pad was new from the factory. Empirically it is found that the break thickness decreases linearly with time, so that $X(t) = \theta t$. Furthermore, a break-pad is considered to have failed if $X(t)$ grows beyond the failure limit L. Now suppose that θ has a normal distribution with mean μ and standard deviation σ. Let T be the time to failure of a break-pad. Now if we wish to determine the time to failure distribution, we have for $t > 0$:

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leq t) \\ &= \mathbb{P}(X(t) \geq L) \\ &= \mathbb{P}(\theta t \geq L) \\ &= \mathbb{P}(\theta \geq L/t) \\ &= \mathbb{P}\left(\frac{\theta - \mu}{\sigma} \geq \frac{L/t - \mu}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{L/t - \mu}{\sigma}\right), \end{aligned} \tag{4.1}$$

where $\Phi(\cdot)$ is the standard normal distribution. This distribution is called the Bernstein distribution. We could carry a similar analysis if θ has a different distribution. From this distribution we can, in principle, determine measures such as the mean time to failure. The assumption that θ has a normal distribution is somewhat peculiar however, because there is always a positive probability that $\theta < 0$, and the break-pad will increase in thickness and never fail. This "detail" actually shows that the mean time to failure of a Bernstein distribution is infinite. However, if $\mathbb{P}(\theta < 0)$ is very small, this is often ignored and then the numerical computation of $\int_{t=0}^{M} \Phi\left(\frac{L/t-\mu}{\sigma}\right) dt$ for some large M yields a numerical approximation for $\mathbb{E}[T]$. ◇

4.1.1 Fitting procedure for random coefficient models

Suppose there is data from m different degradation paths and let x_{ij} denote the degradation level of path $i \in \{1, \ldots, m\}$ at time t_{ij}. Data is available in the form shown in Table 4.1. Least squares regression is the most commonly used technique to fit random coefficient models. For each degradation path $i \in \{1, \ldots, m\}$, an estimate for θ, $\hat{\theta}_i$, is found as the solution to the following optimization problem:

$$\hat{\theta}_i = \text{argmin}_\theta \left(\sum_{j=1}^{n} (f(t_{ij}|\theta) - x_{ij})^2 \right). \tag{4.2}$$

This optimization problem is simple to solve if f is linear or affine[2] in θ (regression). If f is not linear or affine in θ, then the usual approach is to find a suitable transformation that makes it affine or linear. These are standard techniques that can be found in books such as Montgomery and Runger (2007). This procedure will yield an estimate $\hat{\theta}_i$ for each $i \in \{1, \ldots, m\}$. From this sample of size m, one can construct an estimator for the mean of θ, $\hat{\mu_\theta}$, by its sample average

$$\hat{\mu_\theta} = \frac{1}{m} \sum_{k=1}^{m} \hat{\theta}_k, \tag{4.3}$$

and its variance by

$$S_\theta^2 = \frac{1}{m-1} \sum_{k=1}^{m} (\hat{\theta}_k - \hat{\mu_\theta})^2. \tag{4.4}$$

These moment estimators can then be used to fit a distribution on the vector θ.

Table 4.1: Data from of past degradation paths

time item 1	Item 1	time item 2	item 2	...	time item m	item m
$t_{1,1}$	$x_{1,1}$	$t_{2,1}$	$x_{2,1}$	...	$t_{m,1}$	$x_{m,1}$
$t_{1,2}$	$x_{1,2}$	$t_{2,2}$	$x_{2,2}$	...	$t_{m,2}$	$x_{m,2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
$t_{1,n}$	$x_{1,n}$	$t_{2,n}$	$x_{2,n}$	...	$t_{m,n}$	$x_{m,n}$

Example 4.2. Reconsider the break-pads from Example 4.1. Suppose that the degradation paths of six units are given as shown in Table 4.2. These degradation data are visualized in Figure 4.4. We will suppose that $f(t|\theta) = \theta t$. Note that this is different from a usual linear regression model because there is no intercept.[3] The lack of an intercept is equivalent to the assumption that there is no degradation when a break-pad is new. Running a least squares fit for each of these units gives the fits shown in Figure 4.5. The estimators for θ are shown in Table 4.3. The sample mean, $\hat{\mu_\theta}$, is $8.91 \cdot 10^{-5}$ m and the sample standard deviation S_θ is $1.75 \cdot 10^{-5}$ m. We can use these estimates

[2] A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be affine if it can be written as $f(x) = \sum_{i=1}^{n} a_i x_i + b$.

[3] Some care must be taken with computer software here. Most regression packages assume by default that there is an intercept.

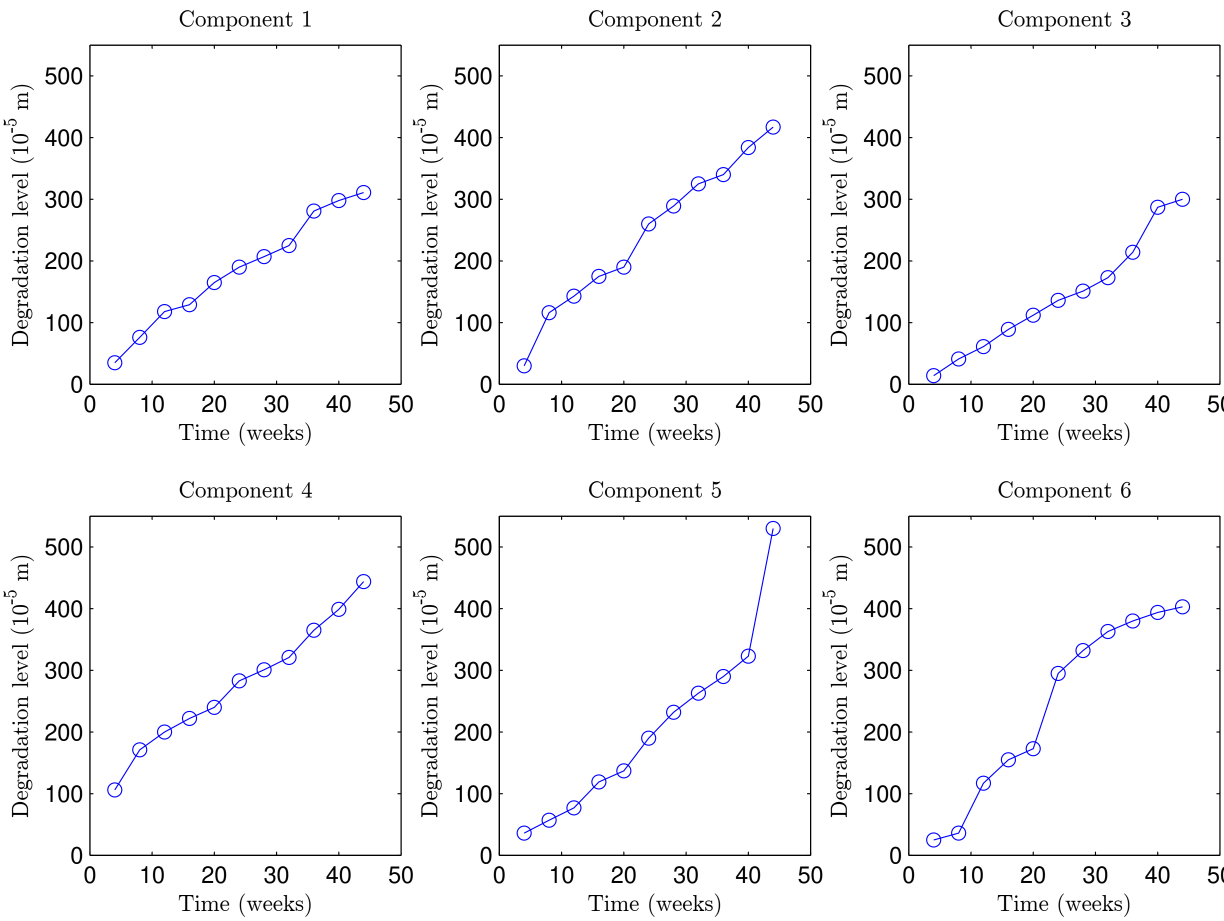


Figure 4.4: Degradation paths from data shown in Table 4.2.

Table 4.2: Example degradation data of six break-pads.

	Wear of breakpad in 10^{-5} m					
Time (in weeks)	1	2	3	4	5	6
4	35	30	14	106	36	25
8	76	116	41	171	57	36
12	118	143	61	200	77	117
16	129	175	89	222	119	155
20	165	190	112	240	137	173
24	190	260	136	283	190	295
28	207	289	151	301	232	332
32	225	325	173	321	263	363
36	281	340	214	365	290	380
40	298	384	287	399	323	394
44	311	417	300	444	530	403

to fit a normal distribution on θ to obtain the Bernstein model from Example 4.1. Under such an estimate, $\mathbb{P}(\theta < 0) = 1.8 \cdot 10^{-7}$ is indeed negligible. $\diamond$

Table 4.3: Estimations of θ for Example 4.2.

Component	1	2	3	4	5	6
$\hat{\theta}$	7.51	9.89	6.20	10.71	8.93	10.24

Remark 4.1. Regression analysis of degradation paths can be a useful tool to find an appropriate transformation that will "linearize" the degradation data. Therefore, is it useful to do some exploratory regression analysis on degradation data before fitting more sophisticated models.

4.2. Gamma process

A Gamma process with shape parameter α and scale parameter β is a process such that

1. $X(0) = 0$

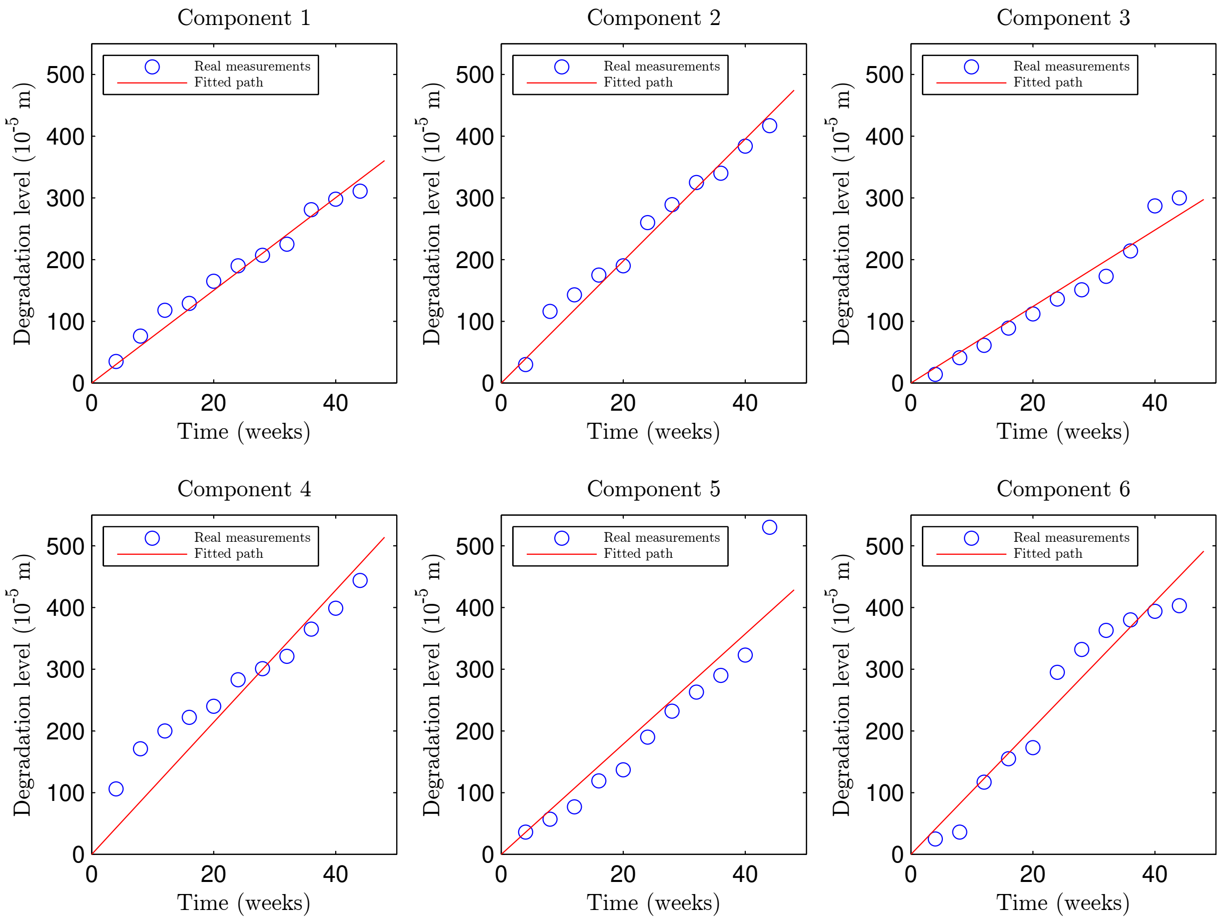


Figure 4.5: Random coefficient models as fitted by least squares estimates for all components.

2. $X(t)$ has independent increments, that is, if $0 \leq t_1 < t_2 < t_3$, then $X(t_3) - X(t_2)$ is independent of $X(t_2) - X(t_1)$.

3. $X(t)$ has stationary Gamma distributed increments, that is for $0 \leq t_1 < t_2$, $X(t_2) - X(t_1)$ has a Gamma distribution with shape $\alpha(t_1 - t_2)$ and scale β.

From this definition we may infer some interesting and useful properties of the Gamma process:

1. The distribution of an increment depends only on the length of that increment. For example $X(8) - X(5)$ has the same distribution as $X(5) - X(2)$, and in general for $0 \leq t_2 < t_2$ and $s > 0$, $X(t_1 + s) - X(t_1)$ has the same distribution as $X(t_2 + s) - X(t_2)$.

2. $X(t)$ is non-decreasing over time. Indeed, for any $t, \varepsilon > 0$, $X(t+\varepsilon) = X(t) + (X(t+\varepsilon) - X(t))$, and since $X(t+\varepsilon) - X(t)$ has a Gamma distribution, it is non-negative. Therefore $X(t+\varepsilon) \geq$

$X(t)$. This property is also intuitive for degradation processes: The condition of equipment should only deteriorate over time, not ameliorate.

3. $X(t)$ is a Markov process, that is, it has the memoryless property. This can easily be verified by observing that for $0 \le t_1 < t_2$:

$$\begin{aligned} \mathbb{P}(X(t_2) \le x \mid (X(t))_{t \le t_1}) &= \mathbb{P}(X(t_1) + (X(t_2) - X(t_1)) \le x \mid (X(t))_{t \le t_1}) \\ &= \mathbb{P}(X(t_1) + (X(t_2) - X(t_1)) \le x \mid X(t_1)). \end{aligned} \tag{4.5}$$

4. The definition of the Gamma process is consistent with known properties of the Gamma distribution. In particular, for $0 \le t_1 < t_2$ we have from the definition that $X(t_2)$ has a Gamma distribution with shape αt_2 and scale β. But we also have that $X(t_2) = (X(t_2) - X(t_1)) + (X(t_1) - X(0))$ so that we must also have shape $\alpha(t_2 - t_1) + \alpha(t_1 - 0) = \alpha t_2$ and scale β. This works out nicely.

5. The Gamma process and its increments have constant variance-to-mean ratios. $\mathbb{E}[X(t)] = \frac{\alpha}{\beta} t$ and $\mathbf{Var}[X(t)] = \frac{\alpha}{\beta^2} t$ so that for all $t \ge 0$

$$\frac{\mathbf{Var}[X(t)]}{\mathbb{E}[X(t)]} = \frac{\frac{\alpha}{\beta^2} t}{\frac{\alpha}{\beta} t} = \beta^{-1}. \tag{4.6}$$

This property also holds for increments, i.e., for any $0 \le t_1 < t_2$, $\frac{\mathbf{Var}[X(t_2) - X(t_1)]}{\mathbb{E}[X(t_2) - X(t_1)]} = \beta^{-1}$.

Another property of the Gamma process that is more subtle, and does not follow immediately from the results above, is that a sample degradation path of a Gamma process is almost nowhere continuous. Despite this, it is common to plot such paths as if they are continuous. Figure 4.6 shows four sample paths of the same process.

An alternative parameterization of the Gamma process is to let $\mu = \alpha/\beta$ and $\sigma = \sqrt{\alpha}/\beta$. This alternative parameterization is easier to interpret because μ and σ^2 represent the increase in mean and variance respectively of $X(t)$ per unit time, i.e.,

$$\mathbb{E}[X(t)] = \mu t \quad \text{and} \quad \mathbf{Var}[X(t)] = \sigma^2 t. \tag{4.7}$$

The original parameters can be recovered from μ and σ as

$$\alpha = \frac{\mu^2}{\sigma^2} \quad \text{and} \quad \beta = \frac{\mu}{\sigma^2}. \tag{4.8}$$

The Gamma process is suitable to model random degradation that is approximately linear in time. If the original degradation process is not linear, it can often be linearized by applying

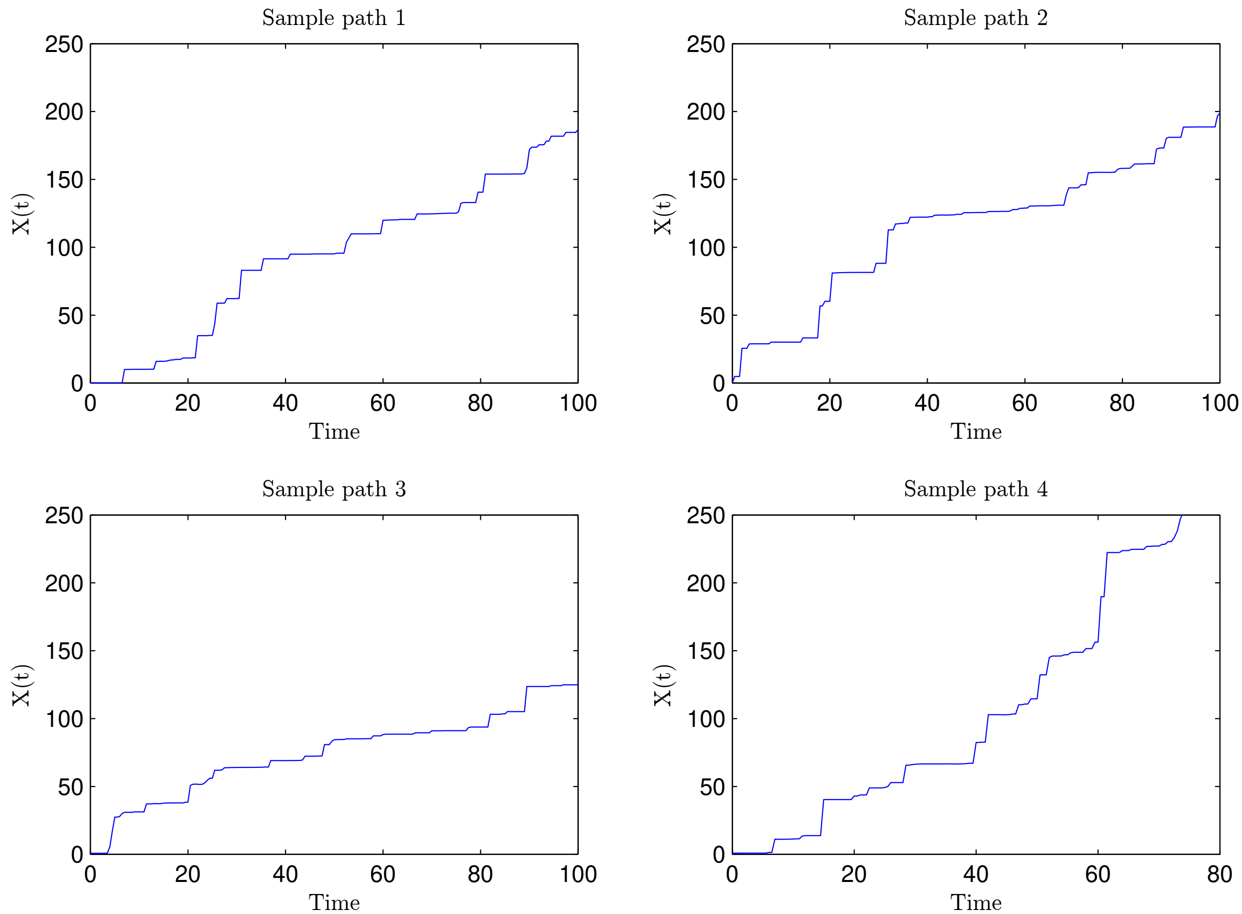


Figure 4.6: Four sample paths of Gamma degradation processes generated by identical processes

appropriate transformations. For example, many mechanical degradation processes are roughly exponential in time, but after a logarithmic transformation, such processes are also linear. Otherwise there exist many generalizations of the Gamma process that allow for modeling more involved non-linear degradation patterns. We refer the reader to van Noortwijk (2009) for an overview of this.

4.2.1 Fitting procedure for the Gamma process

We will find a fit for the Gamma process by estimating the (alternative) parameters μ and σ because they are easier to interpret. The original parameters α and β can then be found via (4.8). We will use the method of moments. Another commonly used method is maximum likelihood estimation, see van Noortwijk (2009). Data is assumed to be available in the same form as explained in Section 4.1.1 and shown in Table 4.1. For convenience assume that $t_{i,1} > 0$ and define $t_{i,0} = 0$ and $x_{i,0} = 0$. Thus, the first measurement is taken after the initial use of an item, and there is no initial degradation[4]. Let $\Delta x_{ij} = x_{i,j} - x_{i,j-1}$ and $\Delta t_{ij} = t_{i,j} - t_{i,j-1}$ for $j \in \{1, \ldots, n\}$. We assume that the times at which data are collected are independent of the degradation levels.

Now suppose that $\Delta t_{ij} = t$ is constant for all measurements. Then Δx_{ij} are i.i.d. samples of a Gamma distribution with mean μt and variance $\sigma^2 t$. Unbiased estimators for μ and σ^2 are therefore given by correcting the sample average and variance for the time length t:

$$\hat{\mu} = \frac{1}{mnt} \sum_{i=1}^{m} \sum_{j=1}^{n} \Delta x_{ij}, \qquad \hat{\sigma^2} = \frac{1}{t} \frac{1}{mn-1} \sum_{i=1}^{m} \sum_{j=1}^{n} (\Delta x_{ij} - \hat{\mu} t)^2. \tag{4.9}$$

When t_{ij} is not the same for all measurements, some care needs to be taken. The following theorem shows how to estimate μ and σ^2 in this case.

Theorem 4.1. *Unbiased estimators for μ and σ^2 are given by*

$$\hat{\mu} = \frac{\sum_{i=1}^m \sum_{j=1}^n \Delta x_{ij}}{\sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij}} \quad \textit{and} \quad \hat{\sigma^2} = \frac{\sum_{i=1}^m \sum_{j=1}^n (\Delta x_{ij} - \hat{\mu} \Delta t_{ij})^2}{\sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij} - \frac{1}{\sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij}} \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2}. \tag{4.10}$$

Proof. Let X_{ij} denote the random variable of which x_{ij} is a realization. Observe that

$$\mathbb{E}\left[\sum_{i=1}^m \sum_{j=1}^n \Delta X_{ij}\right] = \sum_{i=1}^m \sum_{j=1}^n \mu \Delta t_{ij}. \tag{4.11}$$

Rearrangement proves that $\hat{\mu}$ is unbiased.

Next we will compute the expectation of the numerator of $\hat{\sigma^2}$ to show that the denominator of $\hat{\sigma^2}$ is correct:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^m \sum_{j=1}^n \left(\Delta X_{ij} - \frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \Delta t_{ij}\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^m \sum_{j=1}^n \left((\Delta X_{ij} - \mu \Delta t_{ij}) - \left(\frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \Delta t_{ij} - \mu \Delta t_{ij}\right)\right)^2\right] \\
&= \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[(\Delta X_{ij} - \mu \Delta t_{ij})^2\right] \\
&\quad - 2 \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[(\Delta X_{ij} - \mu \Delta t_{ij}) \left(\frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \Delta t_{ij} - \mu \Delta t_{ij}\right)\right] \\
&\quad + \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[\left(\frac{\sum_{k=1}^m \sum_{k=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \Delta t_{ij} - \mu \Delta t_{ij}\right)^2\right].
\end{aligned} \tag{4.12}$$

[4]These assumptions are not difficult to relax, but they make the notation and ideas in this section more transparent.

For convenience, we will work out each of the three terms above separately. The first term simplifies to

$$\textstyle\sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[(\Delta X_{ij} - \mu\Delta t_{ij})^2\right] = \sigma^2 \sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij}. \tag{4.13}$$

The second term simplifies as

$$\begin{aligned}
&\sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[(\Delta X_{ij} - \mu\Delta t_{ij})\left(\frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}}\Delta t_{ij} - \mu\Delta t_{ij}\right)\right] \\
&\quad= \sum_{i=1}^m \sum_{j=1}^n \frac{\Delta t_{ij}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \mathbb{E}\left[(\Delta X_{ij} - \mu\Delta t_{ij})\left(\textstyle\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl} - \mu \sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}\right)\right] \\
&\quad= \frac{1}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^m \sum_{l=1}^n \Delta t_{ij} \mathbf{Cov}(\Delta X_{ij}, \Delta X_{k,l}) \\
&\quad= \frac{1}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij} \mathbf{Var}[\Delta X_{ij}] \\
&\quad= \frac{1}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij} \sigma^2 \Delta t_{ij} \\
&\quad= \frac{\sigma^2}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2,
\end{aligned} \tag{4.14}$$

where the third inequality follows from the independent increments of the Gamma process which implies that $\mathbf{Cov}(\Delta X_{ij}, \Delta X_{kl}) = 0$ whenever $(i,j) \neq (k,l)$. The final term simplifies as follows:

$$\begin{aligned}
&\sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[\left(\frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}}\Delta t_{ij} - \mu\Delta t_{ij}\right)^2\right] \\
&\quad= \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}\left[(\Delta t_{ij})^2 \left(\frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} - \mu\right)^2\right] \\
&\quad= \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2 \mathbf{Var}\left[\frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}}\right] \\
&\quad= \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2 \frac{\mathbf{Var}\left[\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}\right]}{\left(\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}\right)^2} \\
&\quad= \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2 \frac{\sigma^2 \sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}}{\left(\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}\right)^2} \\
&\quad= \frac{\sigma^2}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2
\end{aligned} \tag{4.15}$$

Now combining the three results above we find that

$$\mathbb{E}\left[\sum_{i=1}^m \sum_{j=1}^n \left(\Delta X_{ij} - \frac{\sum_{k=1}^m \sum_{l=1}^n \Delta X_{kl}}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \Delta t_{ij}\right)^2\right]$$

$$= \sigma^2 \sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij} - \frac{2}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sigma^2 \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2 + \frac{1}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sigma^2 \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2$$

$$= \sigma^2 \left(\sum_{i=1}^m \sum_{j=1}^n \Delta t_{ij} - \frac{1}{\sum_{k=1}^m \sum_{l=1}^n \Delta t_{kl}} \sum_{i=1}^m \sum_{j=1}^n (\Delta t_{ij})^2\right) \tag{4.16}$$

Rearranging the last expression completes the proof. □

We remark that there also exist maximum likelihood estimators for the Gamma process, but these are much more involved and require the numerical solution of non-linear equations. Furthermore, the Gamma process can be generalized to allow for non-stationary (but still independent increments) so that non-linear degradation behavior can also be modeled. These non-stationary processes can be fit on data in a similar manner. Interested readers are referred to Zhu (2015) or van Noortwijk (2009).

Example 4.3. Reconsider the data set from Example 4.2 in Table 4.2 depicted in Figure 4.4. This data set has $\Delta t_{ij} = t = 4$ for all i, j. Therefore we can use either (4.9) or Theorem 4.1 to yield $\hat{\mu} = 9.1098$ and $\hat{\sigma^2} = 234.7164$. Using (4.8) we find the estimates for the original parameters $\hat{\alpha} = 0.3536$ and $\hat{\beta} = 0.0388$. We simulated this fitted model four times to verify that the fit is reasonable; compare Figures 4.4 and 4.7.

4.3. Markov degradation processes

This section studies discrete state Markov processes. The Gamma process is also a Markov process in the sense that is has the Markov (memoryless) property, but it is also a continuous state process. This section deals with processes where the degradation state is countable and in many cases even finite. The tools we will use are discrete and continuous time Markov chains as well as semi-Markov processes. Such processes are convenient modeling tools because they lead to computational schemes that are straightforward later on when we will study condition based maintenance policies.

4.3.1 Erlangian degradation

The simplest model of degradation is where there is a finite number of degradation states and a component moves through these states sequentially, sojourning in each one for an exponential amount of time. Each sojourn time also has the same mean and the last state is considered the

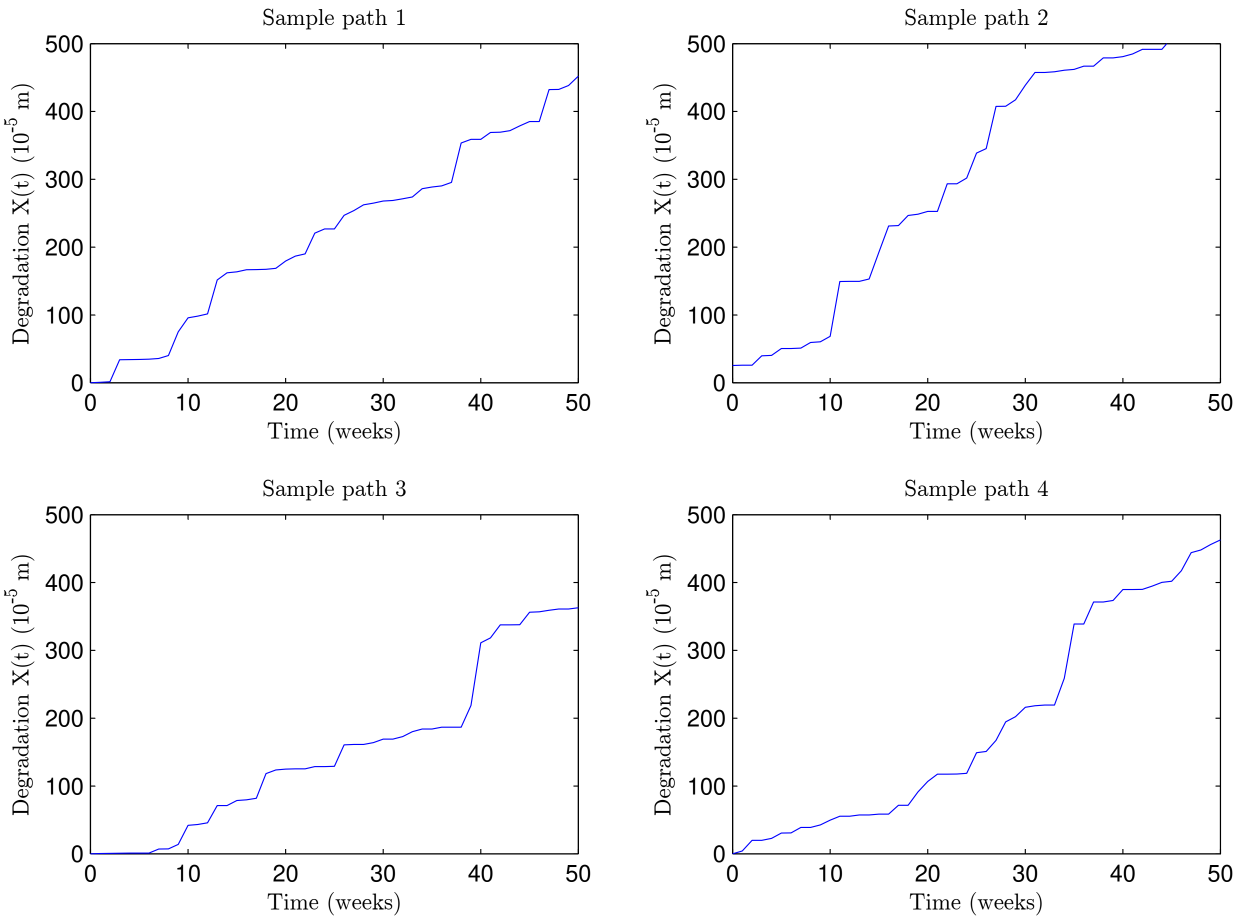


Figure 4.7: Four samples from the Gamma process fitted in Example 4.3.

failed state. The lifetime of such systems is Erlangian (see Section 2.2.3). The state diagram of an Erlang degradation model is shown in Figure 4.8.

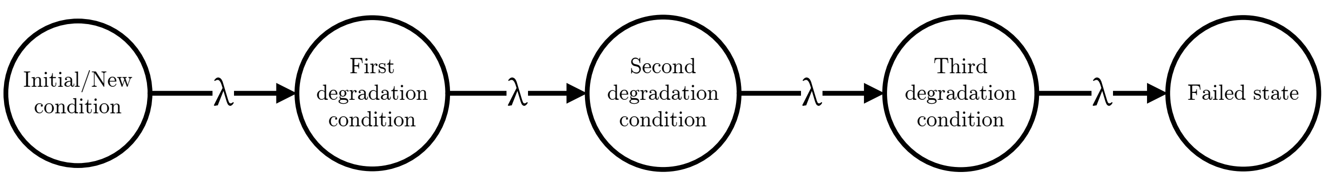


Figure 4.8: Transition diagram of an Erlang degradation model with 5 phases and transition intensity λ.

When the sojourn times in each state have different means, this model is also called the Hypoexponential model (not to be confused with the Hyperexponential model). It is usually difficult to fit these models to real data, but there are some situations where it arises naturally.

In particular, if a system consists of k cold standby redundant components[5], each of which has an exponential lifetime, then the state of this system can be measured by the number of failed components. This leads naturally to this model. Some safety critical electrical systems are set-up this way.

Finally we note that this model is usually formulated in continuous time, but it can also be formulated in discrete time; see the exercises.

4.3.2 Delay time models

A particularly popular class of Markov degradation models are semi-Markov models with only three sequential states as shown in Figure 4.9. Since this is a semi-Markov model, the sojourn time in each state is not necessarily exponential, but can be any distribution with finite mean. The time until a defect occurs is called *time to defect* and the time from then until failure is called *delay time*.

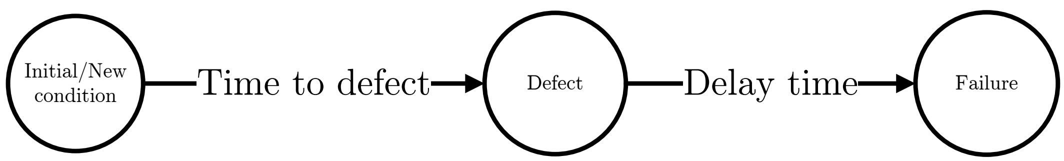


Figure 4.9: Delay time semi-Markov model

Even if the actual degradation measure is a continuous process, it is convenient for engineers and maintenance managers to work with the delay time degradation model because of its simplicity. This is illustrated by the following examples:

- The condition of ball bearing is measured by the amplitude of vibrations. This amplitude is continuous on a certain range. Despite this, it is common practice to divide the whole range of possible vibration amplitudes in three areas: (1) normal operation, (2) defective operation, and (3) failure. If we base maintenance decision only on the area in which the vibration amplitude signal resides, the delay time model occurs naturally.

- The condition of a metal part can be determined by counting the number of cracks and their length in a visual inspection. Here too it is possible to define an area of normal operation, defective operation and failure. The delay time degradation model again becomes natural.

[5] A system has redundant components if is has multiple components of which only one needs to work. As soon as the working component fails, another component takes over. This components is considered redundant before it takes over because it was not actually needed to make the system work. A cold standby component does not degrade unless it is functioning. A warm standby component also degrades while it is still redundant.

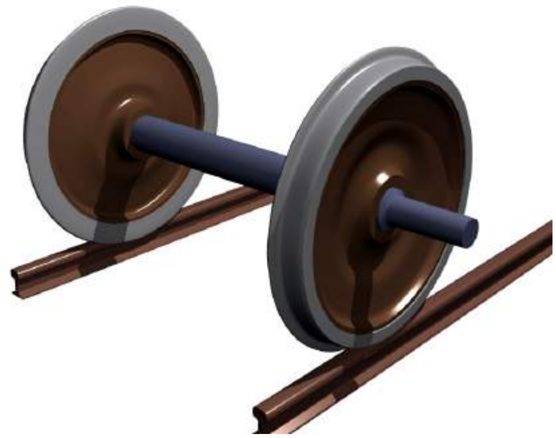
 

Figure 4.10: The ballast on a railway track and the axle of a train bogie

- The concentration of ferrous parts in the lubrication fluid of a metal system with moving parts gives a good indication of the condition. Engineers often define control limits such that if the concentration exceeds these limits, the system is considered defective or failed. Given these control limits, the delay time model again occurs as a natural degradation model.

Furthermore, the delay time model can also be a natural consequence of the physics of failure. This occurs in settings where the degradation processes are initiated by an initial defect. The axle rod of a train is an example where the delay time model is accurate. The axle rod of a train degrades hardly at all until it is somehow damaged by a rock from the ballast hitting it; see Figure 4.10 for an illustration. After a small crack occurs due to such a random event, a crack in the metal will grow according to more or less predictable models of crack growth and propagation from material science. Accidents where the ballast damages the axle rod should occur randomly and uniformly over time. Therefore a usual assumption is that the time to defect is exponentially distributed while the delay time is modeled by a distribution with a much smaller coefficient of variation.

4.3.3 Compound Poisson process

When damage or wear accumulates over time through incidents that occur randomly over time, then the compound Poisson process is a good degradation model. For example, certain metal and ceramic components in structures such as trains and aircraft, only degrade when they are subjected to a load that exceeds some threshold. In materials engineering, it is known that a metal/ceramic deforms when it is subjected to a force at a given temperature (stress). When the stress is sufficiently low, a part will return to its original shape when the stress disappears.

This is called elastic deformation. However, there is a critical stress that depends on the type of metal/ceramic and the component's original form. When the stress applied to a component exceeds this critical level, the part will deform and not return completely to its original form. This is called plastic deformation. Large plastic deformations can hamper system functioning and ultimately lead to component rupture. Therefore, the amount of plastic deformation is usually a good degradation measure. When stress/loads arrive to the system arrive according to a Poisson process and the size of the stress/load has a certain distribution, then the degradation can be modeled by a compound Poisson process.

The compound Poisson process is a process where damage arrives according to a Poisson process with rate λ and the damage done at each arrival is a random variable D. Let $N(t)$ denote the number of load arrivals in $[0, t]$. (Recall that $N(t)$ has a Poisson distribution with mean λt.) Furthermore, let D_i denote the damage corresponding to the i-th damage arrival. Then the degradation process, $X(t)$, is given by

$$X(t) = \sum_{i=1}^{N(t)} D_i. \tag{4.17}$$

D_i is also called the compounding random variable and its distribution is called the compounding distribution. Figure 4.11 shows four sample paths of a compound Poisson degradation process.

Sometimes, we can only observe the degradation at certain points in time $0, \tau, 2\tau, 3\tau, \ldots$. Let $X_t = X(t\tau)$ for $t \in \mathbb{N}_0$. Then we can write

$$X_t = X_{t-1} + Z_t, \tag{4.18}$$

where $(Z_t)_{t\in\mathbb{N}}$ is a sequence of i.i.d. random variables each distributed as $\sum_{i=1}^{N(\tau)} D_i$. The distribution of Z_t depends on λ and the distribution of D_i. For convenience we will now assume that that D_i is a discrete random variable on the integers. If this is not originally so, this can be achieved by discretization and an appropriate choice of measuring unit[6]. Finding the probability mass function of Z_t involves several steps. First let $D^{(n)} = \sum_{i=1}^{n} D_i$. We will now compute the

[6]For example, if damage is measured by deformation and has a Gamma distribution with mean $\mathbb{E}[D] = 2.3 \cdot 10^{-4}$m and standard deviation $\sqrt{\mathbf{Var}[D]} = 1.2 \cdot 10^{-4}$m, then we can choose the micrometer (μm$=10^{-6}$m) as unit and discretize by setting $\mathbb{P}(D = x) \approx F_D((x + 1/2) \cdot 10^{-6}) - F_D((x - 1/2) \cdot 10^{-6})$ for $x \in \mathbb{N}$.

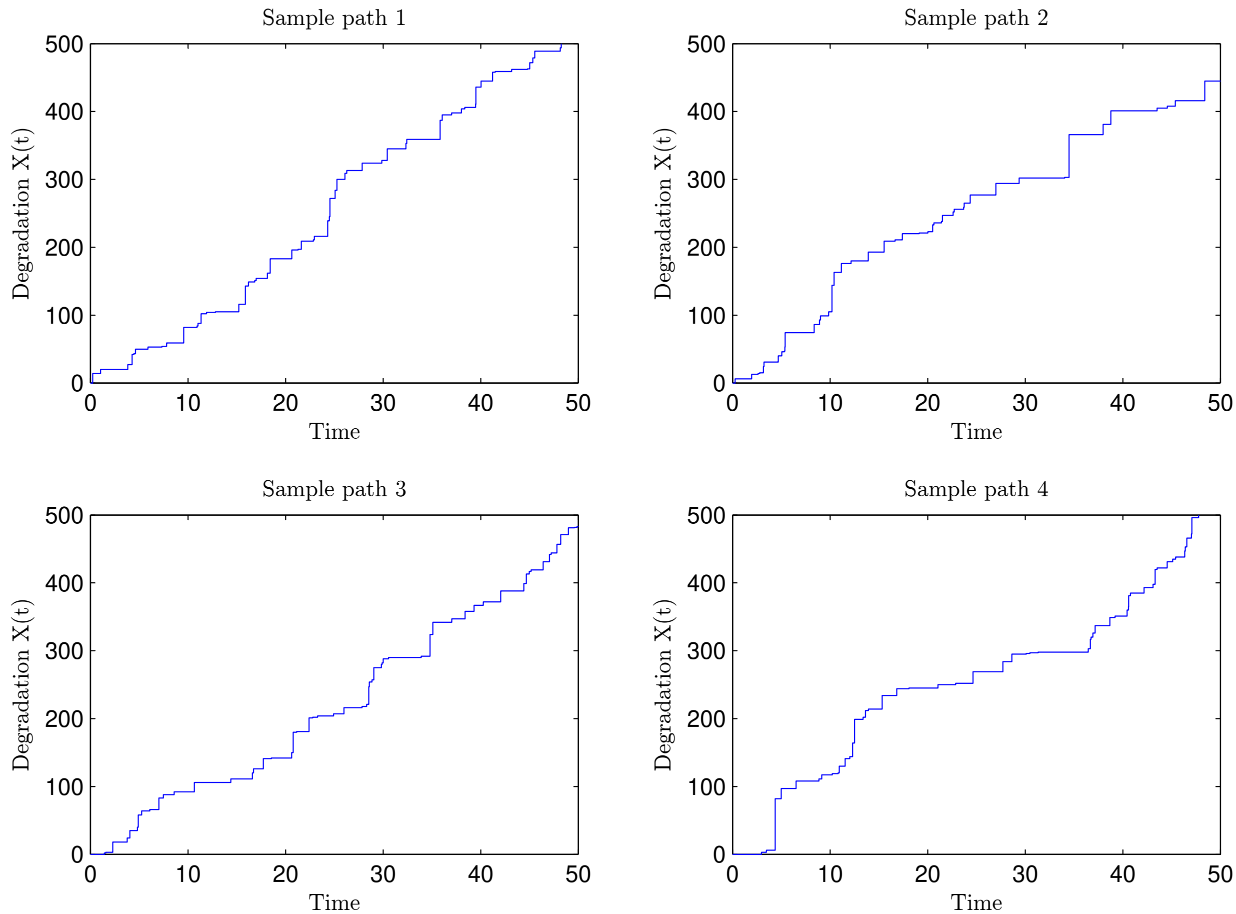


Figure 4.11: Four sample paths of a compound Poisson degradation process where the compounding distribution is geometric.

probability mass function of $D^{(2)}$.

$$\begin{aligned}\mathbb{P}(D^{(2)} = x) &= \mathbb{P}(D_1 + D_2 = x) \\ &= \sum_{y=0}^{x} \mathbb{P}(D_1 + D_2 = x \mid D_2 = y)\mathbb{P}(D_2 = y) \\ &= \sum_{y=0}^{x} \mathbb{P}(D_1 = x - y)\mathbb{P}(D_2 = y). \end{aligned} \tag{4.19}$$

This trick can be applied recursively to find $\mathbb{P}(D^{(n)} = x)$ for any $n \in \mathbb{N}$:

$$\mathbb{P}(D^{(n)} = x) = \sum_{y=0}^{x} \mathbb{P}(D_n = x - y)\mathbb{P}(D^{(n-1)} = y). \tag{4.20}$$

Now we can express the probability mass function of Z_t:

$$\begin{aligned}
\mathbb{P}(Z_t = x) &= \mathbb{P}\left(\textstyle\sum_{i=1}^{N(\tau)} D_i = x\right) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(\textstyle\sum_{i=1}^{N(\tau)} D_i = x \mid N(\tau) = n\right) \mathbb{P}(N(\tau) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(D^{(n)} = x)\mathbb{P}(N(\tau) = n)
\end{aligned} \tag{4.21}$$

The infinite summation in (4.21) needs to be truncated at some large value to actually use this expression for computations. Certain assumptions on the distribution of D yield nice closed form expressions for the distribution of Z_t. If D has a logarithmic distribution with parameter q, then Z_t has a negative binomial distribution with parameters[7]

$$p = 1 - q, \quad r = -\frac{\lambda\tau}{\ln(1-q)}. \tag{4.22}$$

Thus, if D_i has a Logarithmic distribution, then closed form expressions for the distribution of Z_t are available by using (4.22) and Section 2.2.8. In fact, the compound Poisson process with a logarithmic compounding distribution can also be interpreted as a negative binomial process with shape $r = -\frac{\lambda}{\ln(1-q)}$ and scale $p = 1 - q$. The negative binomial process is defined analogously to the Gamma process:

1. $X(0) = 0$
2. $X(t)$ has independent increments (because the Poisson process has independent increments.) That is, if $0 \leq t_1 < t_2 < t_3$, then $X(t_3) - X(t_2)$ is independent of $X(t_2) - X(t_1)$.
3. For $0 \leq t_1 < t_2$, $X(t_2) - X(t_1)$ has a negative binomial distribution with shape $r(t_2 - t_1)$ and scale p.

It is clear by its construction, that the negative binomial process has the same or analogous properties to the Gamma process:

1. The distribution of an increment depends only on the length of that increment.
2. $X(t)$ is non-decreasing over time.
3. $X(t)$ is a Markov process.

[7] A proof of this fact is contained in the appendix. Since this proof uses the Laplace domain, we decided to not put it in the main text.

4. The definition of the negative binomial process is consistent with known properties of the negative binomial distribution. In particular, for $0 \leq t_1 < t_2$ we have from the definition that $X(t_2)$ has negative binomial distribution with shape rt_2 and scale p. But we also have that $X(t_2) = (X(t_2) - X(t_1)) + (X(t_1) - X(0))$ so that we must also have shape $r(t_2 - t_1) + r(t_1 - 0) = rt_2$ and scale p.

5. The negative binomial process and its increments have constant variance-to-mean ratios. $\mathbb{E}[X(t)] = rt\frac{1-p}{p}$ and $\mathbf{Var}[X(t)] = rt\frac{1-p}{p^2}$ so that for all $t \geq 0$

$$\frac{\mathbf{Var}[X(t)]}{\mathbb{E}[X(t)]} = \frac{1}{p} \geq 1. \tag{4.23}$$

This property also holds for increments, i.e., for any $0 \leq t_1 < t_2$, $\frac{\mathbf{Var}[X(t_2)-X(t_1)]}{\mathbb{E}[X(t_2)-X(t_1)]} = p^{-1}$.

4.3.4 Approximating a Gamma process by a compound Poisson process

The compound Poisson process with logarithmic compounding distribution can also be interpreted as a negative binomial process that shares many characteristics with the Gamma process; see Section 4.3.3. Just like the geometric distribution is the discrete analog of the exponential distribution, the negative binomial distribution is the discrete analog of the Gamma distribution. If the mean of a Gamma distribution is sufficiently large, a negative binomial distribution with the same mean and standard deviation usually provides a good fit as shown in Figure 4.12. Such a discrete fit is convenient in many computational procedures where continuous distributions would have to be discretized anyway.

To approximate a Gamma process by a negative binomial process, note that the negative binomial process can be alternatively parameterized by $\mu = r\frac{1-p}{p}$ and $\sigma^2 = r\frac{1-p}{p^2}$ too. These parameters have the same interpretation as under the Gamma process, and can be estimated from data as shown in Theorem 4.1. Estimators for r and p can then be obtained from the identities

$$p = \frac{\mu}{\sigma^2}, \quad r = \frac{p}{1-p}\mu = \frac{\mu^2}{\sigma^2 - \mu}. \tag{4.24}$$

Example 4.4. Reconsider the break-pads from Example 4.2 and the data from Table 4.2. If we fit a negative binomial process on these data using Theorem 4.1 and (4.24), we find the estimators $\hat{r} = 0.36785$, $\hat{p} = 0.03881$. Using (4.22), this is equivalent to a compound Poisson process with arrival intensity $\hat{\lambda} = -\hat{r}\ln(\hat{p}) = 1.195$ and logarithmic compounding distribution with parameter $\hat{q} = 1 - \hat{p} = 0.9612$. Figure 4.13 shows four samples of this fitted compound Poisson process. These should be compared to the original data in Figure 4.4 and sample paths of the fitted Gamma process in Figure 4.7.

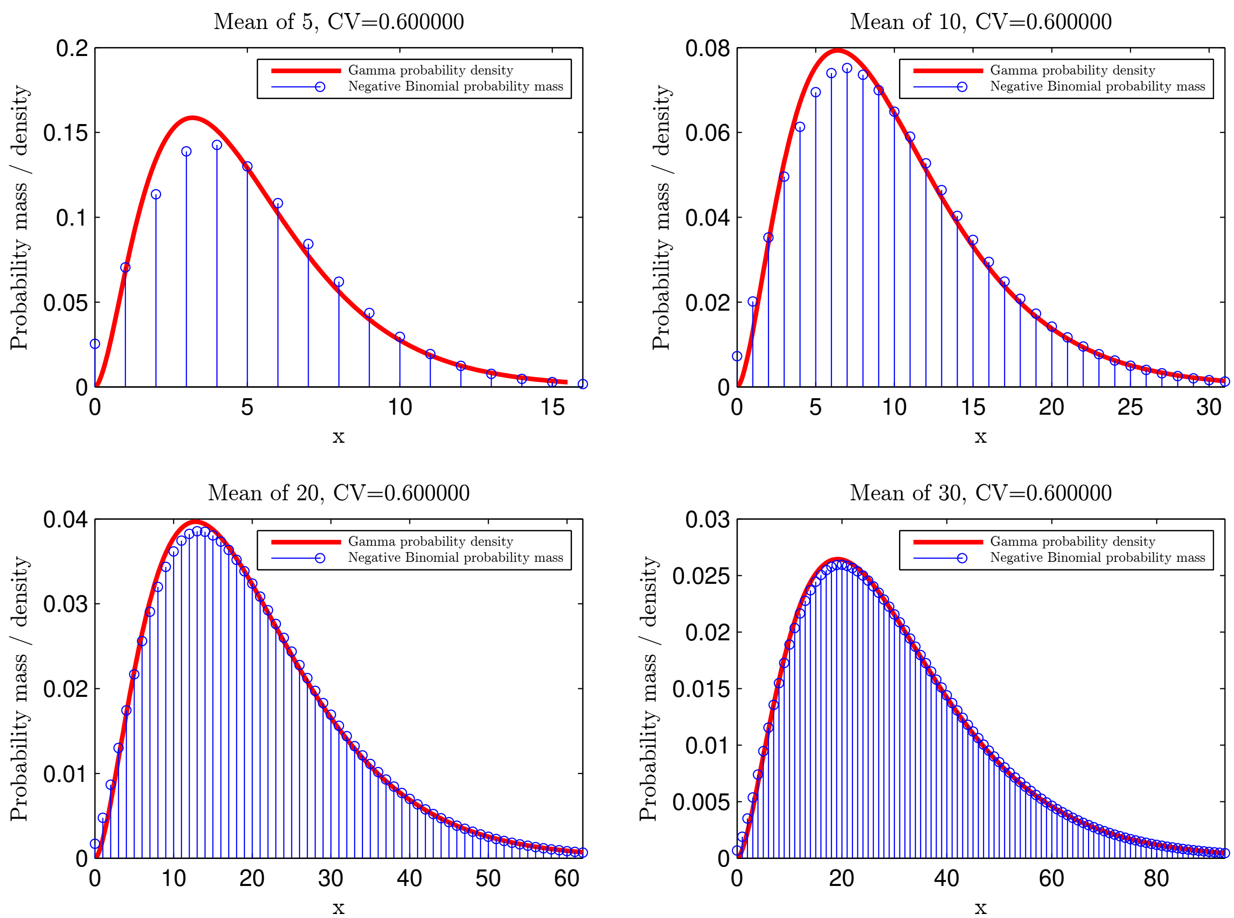


Figure 4.12: Four comparisons of Gamma densities and Negative binomial probability mass function all with a coefficient of variation of 0.6. ($CV = \sqrt{\mathbf{Var}[X]}/\mathbb{E}[X] = 0.6$)

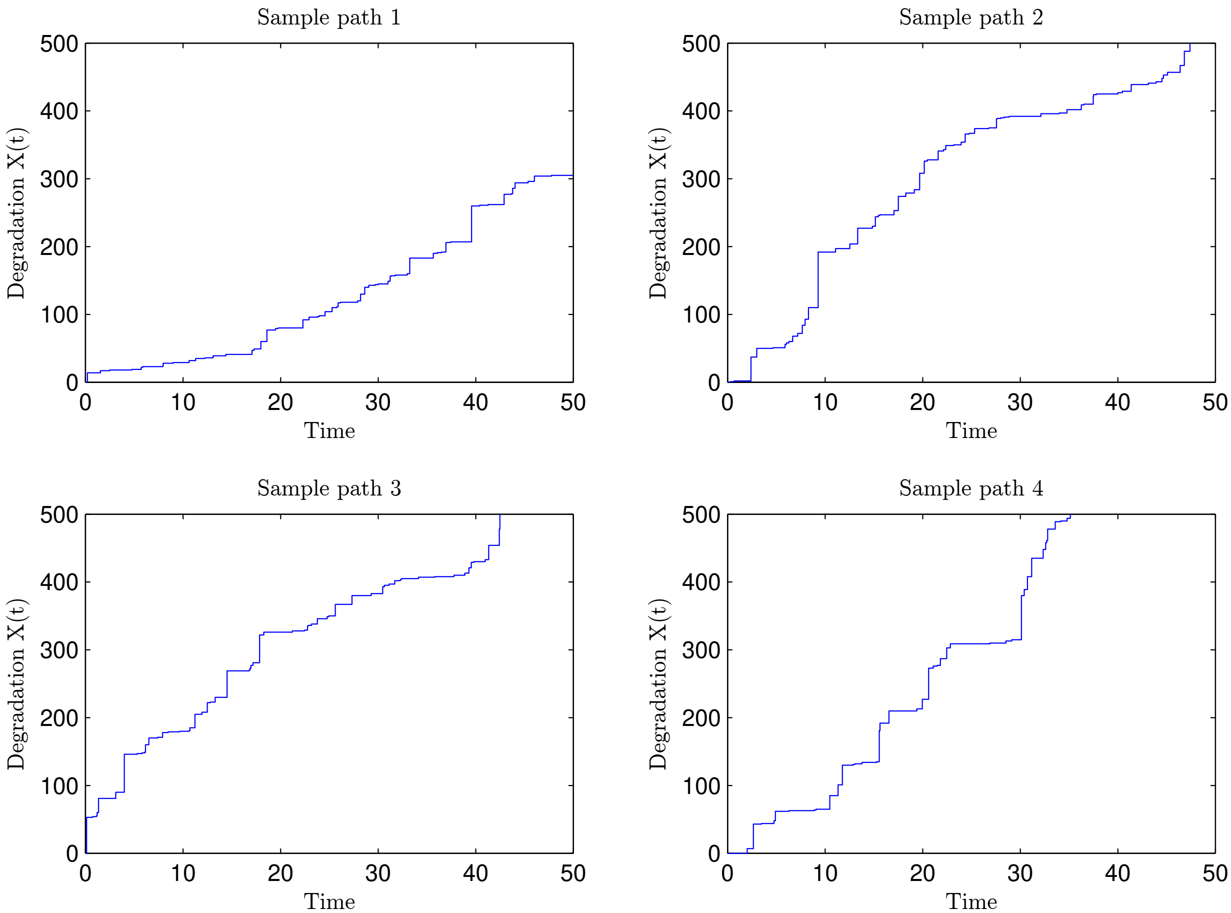


Figure 4.13: Four sample paths of a compound Poisson process with logarithmic compounding fit on the data in Table 4.2

4.4. Non-stationary processes

THIS SECTION IS STILL TO DO AND NOT MATERIAL FOR AMSL

4.5. Exercises

Exercise 4.1. *(Random coefficient model; level: standard; use of computer of graphical calculator permitted for this exercise.)*

One sample degradation path for the degradation of diffusion controlled ageing in the concrete of a nuclear reactor is given in the Table 4.4 below. (We omit units purposefully.)

Table 4.4: Data of diffusion controlled ageing of the concrete in a nuclear reactor

Time	Degradation level
0	0
1	6
2	11
3	13
4	15
5	16
6	18
7	18
8	19
9	21
10	23
11	25
12	24
13	27
14	27
15	29

In order to understand the degradation character, we would like to under stand the functional form of the degradation path. Let $x(t)$ denote the degradation level at time t.

(a) Suppose that the functional form is given by $x(t) = at$. Give a point estimate for a based on this degradation path using least squares regression.

(b) Suppose that the functional form is given by $x(t) = at + b$. Give a point estimates for a and b based on this degradation path using least squares regression.

(c) It is known from literature (Ellingwood and Mori (1993)) that this type of degradation grows as the square root, i.e., $x(t) = a\sqrt{t}$. Use least squares regression to estimate a for this degradation model.

(d) A more general functional form is $x(t) = at^b$. Use least squares regression to estimate a and b. (Hint: a logarithmic transformation may be useful to accomplish this.) Is your estimate for b close to the $\frac{1}{2}$ suggested in the literature?

(e) Suppose that in fact $x(t) = a\sqrt{t}$ and a has a normal distribution with mean μ and standard deviation σ and that failure occurs when $x(t)$ exceeds L for the first time. Let T be the lifetime of the concrete, i.e., $T = \inf\{t \mid x(t) \geq L\}$. Express $\mathbb{P}(T \leq t)$ in terms of μ, σ, L and the standard normal random variable. Compute $\mathbb{P}(T < t)$ for $\mu = 7$, $\sigma = 1$, $L = 12$ and $t = 2$.

Exercise 4.2. *(Random coefficient model; level: standard; use of computer or graphical calculator admitted.)*

Consider the data set below in Table 4.5 about the degradation of concrete subject to sulphate attack.

Table 4.5: Concrete sulphate attack degradation data.

Time	Degradation level
0	1
1	0
2	0
3	2
4	4
5	4
6	5
7	7
8	8
9	11
10	14
11	18
12	21
13	25
14	28
15	34

(a) Use this degradation path to fit a linear model (no intercept) and an affine model (linear model with intercept) using least squares regression. What are the values of the slopes and intercept?

(b) According to Ellingwood and Mori (1993), degradation of this kind should have a quadratic form. Fit a quadratic model without intercept and linear term.

(c) Verify whether the quadratic form is appropriate by fitting a general power model of the form $x(t) = at^b$. What is your estimate for b?

(d) Do you expect the Gamma process to provide a good fit for this data? Why (not)?

Exercise 4.3. *(Gamma process and fitting procedure)*

An engineer plans to study the degradation of a roller bearing. He has an experimental setup where he can monitor the degradation signal in continuous time. Storing the entire degradation signal however is not feasible due to computer memory considerations. He comes up with the following idea: Each time the degradation signal exceeds $2n$, $n \in \mathbb{N}$ for the first time, both the degradation signal and the time this happens are recorded. (Note that this will lead to a data set of the form shown in Table 4.1.) The engineer then plans to use this data to estimate μ and σ using Theorem 4.1. Do you believe this combination of sampling plan and estimation procedure is a good idea? Why (not)? If you believe the plan is bad, can you suggest a better alternative?

Exercise 4.4. *(Markov and Erlangian models)*

The navigation unit on a transatlantic ship is an essential component for safety. A navigation unit only needs one antenna, but it usually has several redundant antennas for safety considerations. Suppose a system has three antenna: the best one has an exponential lifetime with failure rate λ and the other two also have an exponential lifetime, but with a higher failure rate of θ.

(a) Suppose that the ship uses only one antenna at any time, starting with the best one. Suppose that the lifetime of an antenna starts only after it is put into use. Make a Markov model for the degradation of the antenna unit of the navigation system. What is the mean time to failure and variance of the time to failure?

(b) Suppose now that the navigation system actually uses all three antennas (so that the lifetime of each antenna starts immediately) but that the navigation system will remain working as long as at least one antenna is working. Make a Markov model for the degradation of this system.

Exercise 4.5. *(Erlangian degradation)*

We consider an Erlangian degradation process with 6 states numbered $0, 1, \ldots, 5$ where 5 denotes the failed state. The rate out of each state to the next degradation state is λ. However, this system is only inspected every τ time units. Let p_{ij} denote the probability of jumping from state i to state j between inspections.

(a) Give a general expression of p_{ij} for $i, j \in \{0, 1 \ldots, 5\}$ as a function λ and τ

(b) Compute p_{0j} for $j = 0, \ldots, 5$ if $\lambda = 2$ and $\tau = \frac{1}{2}$.

Exercise 4.6. *(Compound Poisson process; level: advanced)*

A particular wing section of an aircraft is loaded whenever turbulence in a specific frequency range occurs. It is known from data that such turbulences occur 10.0 times per year according to a Poisson process. The stress experienced at the wing section (X_i) is measured in kN and has an Erlang distribution with shape $k = 2$ and scale $\lambda = 1$ kN. The wing section only degrades when the stress exceeds 5.00 kN and the strain/plastic deformtion/degradation is 0.1 mm for every 0.2 kN of stress beyond 5.00 kN.

(a) Argue that the amount of degradation deformation in mm at a turbulence incident is $\max(0, X_i - 5.0)/(0.2/0.1) = (X_i - 5)^+/2$, with $x^+ = \max(0, x)$.

(b) What fraction of turbulence incidents will lead to actual damage? What is the effective arrival rate if we use the Compound Poisson distribution to model degradation?

(c) Suppose that the aircraft can fail catastrophically when the plastic deformation exceeds 6 mm. If a new aircraft makes a flight and encounters turbulence, what is the probability of a catastrophe?

(d) Suppose that a turbulence incident leads to degradation. Let D_i denote the amount of degradation (and recall that this is conditional on degradation happening in the first place). Find an expression for $\mathbb{P}(D_i \leq x)$.

Exercise 4.7. *(Gamma process; level: below standard)*

Suppose that the degradation of a component occurs according to a Gamma process with shape $\alpha = 1.2$ and scale $\beta = 2$. A component fails when the degradation exceeds 15. Compute the probability that the components is still working at time 5.

Exercise 4.8. *(Delay time model; level: standard)*

Consider a component whose degradation can be modeled by an exponential time to defect time with mean 1, and a delay time that is uniform between 0.4 and 0.6. Let T denote the lifetime of this components. Use conditioning to compute the distribution function of T.

Exercise 4.9. The degradation data of two pantographs are shown in Table 4.6. The maintenance engineer would like to fit either a gamma process or a compound Poisson process with logarithmic compounding on this data.

Table 4.6: Pantograph degradation data of two parts

time 1	degradation 1	time 2	degradation 2
0	0	0	0
13	5	16	7
29	12	28	16
39	18	39	20
55	26	53	27
64	31	70	32
75	34	-	-
94	40	-	-
106	47	-	-

(a) Explain why Theorem 4.1 cannot be applied directly. (Hint: What is n and what is m for these data?)

(b) The problem highlighted under (a) can be easily fixed by replacing n by n_m in Theorem 4.1. What is the interpretation of n_m and what are n_1 and n_2 for the pantograph data set?

(c) Fit a Gamma process on the pantograph degradation data. What do you find for $\hat{\mu}$, $\hat{\sigma^2}$, $\hat{\alpha}$, and $\hat{\beta}$?

(d) Fit a compound Poisson process with logarithmic compounding on the pantograph data. If you cannot, explain why this is not possible.

(e) Suppose that $\hat{\mu} = 5$ and $\hat{\sigma^2} = 17$. Fit a compound Poisson process based on these estimators.

4.5.1 Solutions

Solution 4.1. (a) $\hat{a} = 2.2089$.

(b) $\hat{a} = 1.6147$, $\hat{b} = 6.1397$.

(c) $\hat{a} = 7.2315$.

(d) $x(t) = at^b$ implies also that $\ln(x(t)) = \ln(at^b) = \ln(a) + \ln(t^b) = \ln(a) + b\ln(t)$. However, $\ln(0) = -\infty$ so that we must omit the first observation from this fitting model. We find the estimates $\hat{b} = 0.5276$ and $\widehat{\ln(a)} = 1.9184$ so that $\hat{a} = \exp(\widehat{\ln(a)}) = 6.8103$.

(e) $\mathbb{P}(T \leq t) = \mathbb{P}(a\sqrt{t} \geq L) = \mathbb{P}(a \geq L/\sqrt{t}) = 1 - \Phi((L/\sqrt{t} - \mu)/\sigma)$. $\mathbb{P}(T < 2) = 1 - \Phi(12/\sqrt{2} - 7) = 0.0687$

Solution 4.2. (a) Linear: $x(t) = at$, $\hat{a} = 1.6944$. Affine: $x(t) = at + b$, $\hat{a} = 2.1647$, $\hat{b} = -4.8603$.

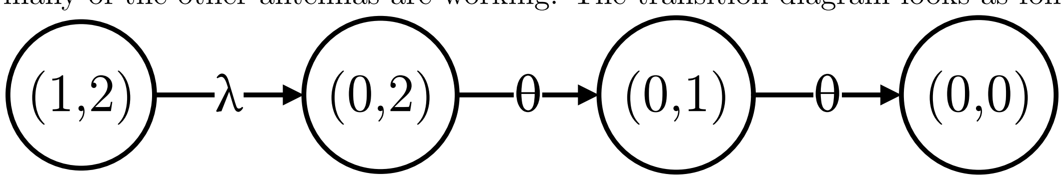
(b) Quadratic: $x(t) = at^2$, $\hat{a} = 0.1462$.

(c) Power law: $x(t) = at^b$; omit first three observation because of zeros; $\hat{a} = 0.2717$, $\hat{b} = 1.7308$.

(d) No. The mean of a Gamma process grows linearly with time and this degradation process is far from linear.

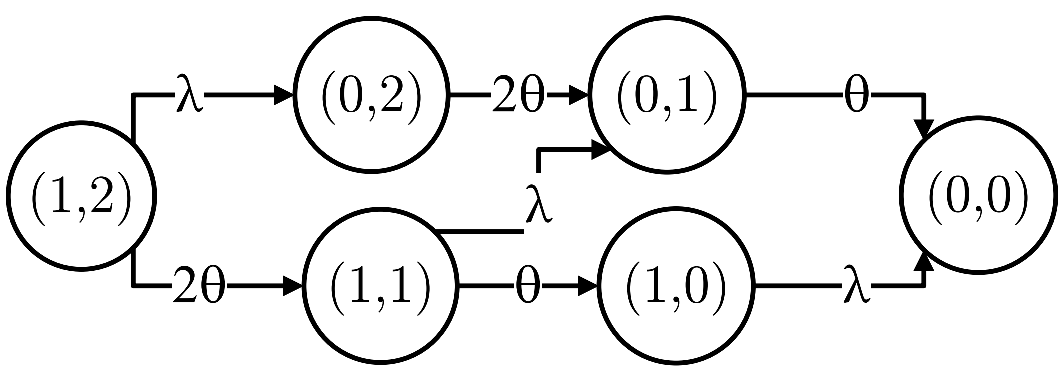
Solution 4.3. This combination of sampling plan and estimation is a bad idea because the time measurements and degradation measurements will not be independent. Therefore the estimation procedure is inappropriate. A better alternative is to measure the degradation signal at equally spaced points in time.

Solution 4.4. (a) The state space for this Markov model is given by $(i, j) \in \{0, 1\} \times \{0, 1, 2\}$ where $i \in \{0, 1\}$ denotes whether the best antenna is working and $j \in \{0, 1, 2\}$ denotes how many of the other antennas are working. The transition diagram looks as follows.



Let T denote the lifetime of the antenna system. We have $\mathbb{E}[T] = 1/\lambda + 2/\theta$ and $\mathbf{Var}[T] = 1/\lambda^2 + 2/\theta^2$.

(b) We use the same state space as the previous subquestion. The transition diagram now looks as follows.



Solution 4.5. (a) Let X denote a Poisson random variable with mean $\lambda\tau$. Then we have $q_x = \mathbb{P}(X = x) = \frac{(\lambda\tau)^x}{x!}\exp(-\lambda\tau)$ and $q_{\geq x} = 1 - \sum_{k<x} q_k$. Now we can write:

$$p_{ij} = \begin{cases} q_{j-i}, & \text{if } j < 5 \text{ and } i \leq j; \\ q_{\geq j-i}, & \text{if } i < j \text{ and } j = 5; \\ 1, & \text{if } i = j = 5; \\ 0, & \text{if } j < i. \end{cases}$$

(b) Use that $q_x = \frac{\lambda\tau}{x} q_{x-1}$ to find quickly $p_{00} = 0.3679$, $p_{0,1} = 0.3679$, $p_{0,2} = 0.1839$, $p_{0,3} = 0.0613$, $p_{0,4} = 0.0153$, and $p_{0,5} = 0.0031$.

Solution 4.6. (a) -

(b) Fraction that leads to damage is $\mathbb{P}(X_i > 5.0) = \sum_{n=0}^{1} \frac{5^n}{n!}\exp(-5) = 6\exp(-5) = 0.0404$. Effective arrival rate of damage incidents per year is therefore $10.0 \cdot 6e^{-5} = 0.4043$ incidents per year.

(c) A plastic deformation of 6 mm occurs when $X_i > 5 + 0.2 \cdot 6/0.1 = 17$ kN. $\mathbb{P}(X_i > 17) = 18e^{-17} = 7.4519 \cdot 10^{-7}$.

(d)

$$\begin{aligned} \mathbb{P}(D_i \leq x) &= \mathbb{P}((X_i - 5)/2 \leq x \mid X_i > 5) \\ &= \frac{\mathbb{P}((X_i - 5)^+/2 \leq x \cap X_i > 5)}{\mathbb{P}(X_i > 5)} \\ &= \frac{\mathbb{P}(5 < X_i \leq 2x + 5)}{\mathbb{P}(X_i > 5)} \\ &= \frac{\mathbb{P}(X_i \leq 2x + 5) - \mathbb{P}(X_i \leq 5)}{\mathbb{P}(X_i > 5)} \\ &= \frac{6e^{-5} - (2x + 6)e^{-2x-5}}{6e^{-5}}. \end{aligned} \tag{4.25}$$

Solution 4.7. Let $X(t)$ denote the degradation level. Then $X(5)$ has a Gamma distribution with $\alpha = 1.2 \cdot 5 = 6$ and $\beta = 2$. Let T be the lifetime of the component. Since α is integer we have: $\mathbb{P}(T > 5) = \mathbb{P}(X(5) > 15) = \sum_{n=0}^{\alpha-1} \frac{(\beta L)^n}{n!}\exp(-\beta L) = \sum_{n=0}^{5} \frac{30^n}{n!}\exp(-30) = 2.2573 \cdot 10^{-8}$.

Solution 4.8. Let $T = X + U$ where X is exponential with mean 1 and U is uniform between 0.4 and 0.6. We have for $t \geq 0.4$ and using the substitution $x = -t - u$:

$$\begin{aligned}
\mathbb{P}(T \leq t) &= \mathbb{P}(X + U \leq t) \\
&= \int_{u=0.4}^{0.6} \mathbb{P}(X + U \leq t \mid U = u) 5du \\
&= \int_{u=0.4}^{0.6} 5\mathbb{P}(X \leq t - u) du \\
&= \int_{u=0.4}^{0.6} 5(1 - e^{-(t-u)}) du \\
&= \int_{u=0.4}^{0.6} 5du - 5 \int_{u=0.4}^{0.6} e^{-(t-u)} du \\
&= [5u]_{u=0.4}^{0.6} + 5 \int_{x=-(t-0.4)}^{-(t-0.6)} e^{x} dx \\
&= 1 + 5[e^{x}]_{x=-(t-0.4)}^{-(t-0.6)} \\
&= 1 + 5(e^{-(t-0.6)} - e^{-(t-0.4)}). \qquad (4.26)
\end{aligned}$$

For $t < 0.4$ a similar derivation can be made.

Solution 4.9. (a) -

(b) n_m is the number of measurements for component m. $n_1 = 8$, $n_2 = 5$.

(c) $\hat{\mu} = 0.448864$, $\hat{\sigma^2} = 0.239513$, $\hat{\alpha} = 0.841201$, $\hat{\beta} = 1.874069$.

(d) This is not possible because the estimated variance to mean ration $\hat{\sigma^2}/\hat{\mu} \approx 0.5 < 1$ and the negative binomial process must have a variance to mean ratio of at least one.

(e) First we fit a negative binomial process and then reinterpret that as a compound Poisson process. We find for the negative binomial process $\hat{p} = 5/17 = 0.2941$ and $\hat{r} = 5^2/(17 - 5) = 2.0833$. Fitting a compound Poisson process we have for the parameter q of the logarithmic compounding distribution $\hat{q} = 1 - \hat{p} = 0.7059$, and for the arrival rate $\hat{\lambda} = -\hat{r}\ln(\hat{p}) = 2.5495$ per time unit.

Chapter 5

Condition Based Maintenance

> "There is only one kind of shock worse than the totally unexpected: the expected for which one has refused to prepare"
>
> Mary Renault

As in Chapter 3, we are still mostly interested in the questions "When should we perform maintenance?" and "What is the cost associated with a certain maintenance policy?". Rather than only using the usage or age of a component to address these questions, we will use information about the condition/degradation of components. Chapter 4 provides models for the condition/degradation level over time of a component as well as fitting procedures. This Chapter has two main sections. Section 5.1 considers delay time degradation models and focusses on optimizing the length of inspection intervals. Section 5.2 studies discrete state Markovian degradation and studies what degradation levels should trigger preventive replacement.

5.1. Maintenance policies for delay time degradation

A component degrades according to a delay time model where the time to defect is denoted by X and the delay time by Y. We assume these random variables have densities and distributions given by $f_X(\cdot), f_Y(\cdot)$ and $F_X(\cdot), F_Y(\cdot)$ respectively. The next three subsection will make different assumptions about X, Y and inspection regimes, and show how to evaluate and optimize these inspection regimes.

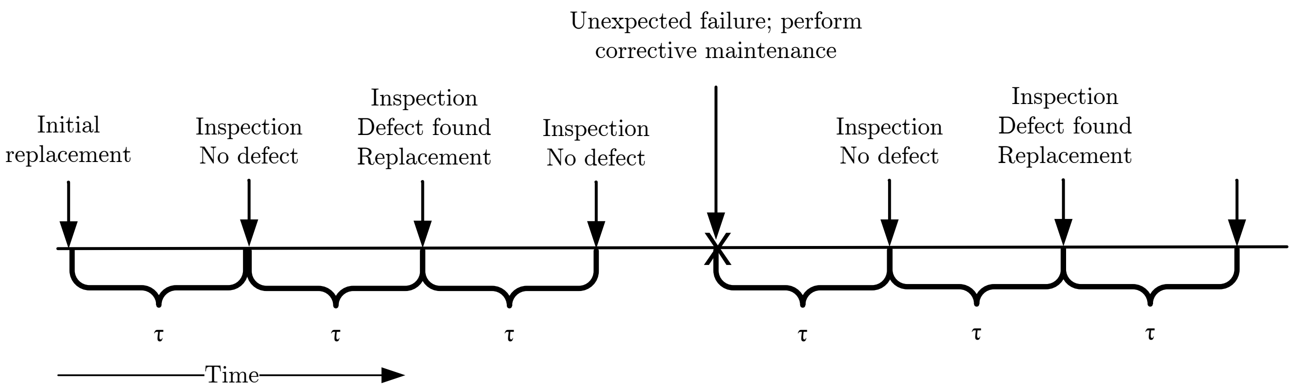


Figure 5.1: Sample path of delay time model subject to periodic inspection and emergency corrective maintence

5.1.1 Exponential time to defect and emergency corrective maintenance

Assume that X has an exponential distribution with mean μ^{-1} and that we inspect the component at times $\tau, 2\tau, 3\tau \dots$ after it has been replaced. If a defect is found at one of these inspection epochs, the component is replaced. We assume that the component will be replaced correctively immediately after failure in an emergency procedure. (This requires that the failure is *self announcing*. If an emergency-break-system fails between inspection epochs, it will likely go unnoticed and this assumption will not hold. If the component is the crankshaft of an engine, the failure will be self announcing.) A sample path of this maintenance policy is shown in Figure 5.1. We assume that inspections cost C_i per inspection. These costs include the cost of performing the inspection itself, but they may also include the cost for the down-time of the asset while the system is down for inspection. We will use renewal reward theory to analyze this policy.

Since X has an exponential distribution, a renewal cycle starts at each inspection epoch and at each corrective replacement. The lack of memory of the exponential distribution is crucial here: When no defect is found during an inspection, the component can be considered as good as new. There are three events that can end a renewal cycle: (1) the component fails between inspection epochs and is replaced correctively, (2) a defect is found during inspection and the component is replaced preventively, and (3) no defect is found during an inspection epoch and the components is as good as new. Now to analyze this renewal cycle, we will need the following preliminaries.

Let $T = X + Y$ and note that by the independence of X and Y we have

$$\begin{aligned}\mathbb{P}(T \leq t) = F_T(t) &= \mathbb{P}(X + Y \leq t) \\ &= \int_{x=0}^{t} \mathbb{P}(X + Y \leq t \mid X = x) f_X(x) dx \\ &= \int_{x=0}^{t} F_Y(t - x) f_X(x) dx, \end{aligned} \tag{5.1}$$

and similarly

$$f_T(t) = \frac{d}{dt} F_T(t) = \int_{x=0}^{t} f_Y(t - x) f_X(x) dx. \tag{5.2}$$

The expected cycle length is given by

$$ECL = \mathbb{E}[\min(T, \tau)] = \int_{x=0}^{\tau} t f_T(t) dt + \tau(1 - F_T(\tau)). \tag{5.3}$$

For the expected cycle cost, we distinguish three types of cycles:

1. A component fails before the next inspection epoch and incurs a cost of C_u. This event occurs with probability $\mathbb{P}(T < \tau) = F_T(\tau)$

2. A component is defective at the next inspection epoch and incurs a cost of C_p for the preventive replacement and C_i for the inspection. This event occurs with probability

$$\begin{aligned}\mathbb{P}(X < \tau \cap X + Y > \tau) &= \int_{x=0}^{\tau} \mathbb{P}(X < \tau \cap X + Y > \tau | X = x) f_X(x) dx \\ &= \int_{x=0}^{\tau} \mathbb{P}(Y > \tau - x) f_X(x) dx \\ &= \int_{x=0}^{\tau} (1 - F_Y(\tau - x)) f_X(x) dx. \end{aligned} \tag{5.4}$$

3. A component still functions normally at the next inspection epoch and only incurs the inspection cost C_i. This event occurs with probability $\mathbb{P}(X > \tau) = 1 - F_X(\tau)$.

Combining the results of the analysis above we find

$$ECC = C_u F_T(\tau) + (C_i + C_p) \int_{x=0}^{\tau} (1 - F_Y(\tau - x)) f_X(x) dx + C_i(1 - F_X(\tau)), \tag{5.5}$$

so that the cost-rate is $g(\tau) = ECC/ECL$. The optimal inspection interval τ^* can be found by minimizing $g(\tau)$. Although $g(\tau)$ is not in general convex, it is usually uni-modal such that greedy algorithms and first order conditions can be used to find the optimum.

Example 5.1. Consider a part where the time to defect has an exponential distribution with rate $\lambda_X = 0.6$ and the delay time has an exponential distribution with rate $\lambda_Y = 0.75$. The time unit is years, corrective maintenance costs 1000, preventive maintenance costs 100 and inspections cost 15 EuroS. Now if we let X denote the time to defect, Y the delay time, and $T = X + Y$, then by using (5.2) we find

$$\begin{aligned} f_T(t) &= \int_{x=0}^{t} \lambda_X e^{-\lambda_X x} \lambda_Y e^{-\lambda_Y (t-x)} dx \\ &= \lambda_X \lambda_Y \int_{x=0}^{t} e^{-\lambda_X x} e^{\lambda_Y x} e^{-\lambda_Y t} dx \\ &= \lambda_X \lambda_Y e^{-\lambda_Y t} \int_{x=0}^{t} e^{-(\lambda_X - \lambda_Y) x} dx \\ &= \lambda_X \lambda_Y e^{-\lambda_Y t} \left[\frac{e^{-(\lambda_X - \lambda_Y) x}}{\lambda_Y - \lambda_X} \right]_{x=0}^{t} \\ &= \lambda_X \lambda_Y e^{-\lambda_Y t} \left[\frac{e^{-(\lambda_X - \lambda_Y) t}}{\lambda_Y - \lambda_X} - \frac{1}{\lambda_Y - \lambda_X} \right] \\ &= \frac{\lambda_X \lambda_Y (e^{-\lambda_X t} - e^{-\lambda_Y t})}{\lambda_Y - \lambda_X}. \end{aligned} \tag{5.6}$$

The easiest way to compute $F_T(t)$ is now (try this yourself)

$$F_T(t) = \int_{x=0}^{t} f_T(x) dx = 1 - \frac{\lambda_Y e^{-\lambda_X t} - \lambda_X e^{-\lambda_Y t}}{\lambda_Y - \lambda_X}, \tag{5.7}$$

so that

$$\begin{aligned} ECL &= \int_{t=0}^{\tau} t \frac{\lambda_X \lambda_Y (e^{-\lambda_X t} - e^{-\lambda_Y t})}{\lambda_Y - \lambda_X} dt + \tau \frac{\lambda_Y e^{-\lambda_X \tau} - \lambda_X e^{-\lambda_Y \tau}}{\lambda_Y - \lambda_X} \\ &= \frac{\lambda_X \lambda_Y}{\lambda_Y - \lambda_X} \left[\frac{1 - e^{-\lambda_X \tau}(\lambda_X \tau + 1)}{\lambda_X^2} - \frac{1 - e^{-\lambda_Y \tau}(\lambda_Y \tau + 1)}{\lambda_Y^2} \right] + \tau \frac{\lambda_Y e^{-\lambda_X \tau} - \lambda_X e^{-\lambda_Y \tau}}{\lambda_Y - \lambda_X}, \end{aligned} \tag{5.8}$$

and

$$\begin{aligned} ECC &= 1000 F_T(\tau) + 15 e^{-\lambda_X \tau} + 115 \int_{t=0}^{\tau} \lambda_X e^{-\lambda_X t} e^{-\lambda_Y (\tau - t)} dt \\ &= 1000 F_T(\tau) + 15 e^{-\lambda_X \tau} + 115 \lambda_X e^{-\lambda_Y \tau} \int_{t=0}^{\tau} e^{-(\lambda_X - \lambda_Y) t} dt \\ &= 1000 F_T(\tau) + 15 e^{-\lambda_X \tau} + 115 \lambda_X e^{-\lambda_Y \tau} \left[\frac{e^{-(\lambda_X - \lambda_Y) t}}{\lambda_Y - \lambda_X} \right]_{t=0}^{\tau} \\ &= 1000 F_T(\tau) + 15 e^{-\lambda_X \tau} + 115 \lambda_X \frac{e^{-\lambda_X \tau} - e^{-\lambda_Y \tau}}{\lambda_Y - \lambda_X}. \end{aligned} \tag{5.9}$$

Using these expressions we can compute $g(\tau) = ECC/ECL$ as shown in Figure 5.2. It is now easy to verify that $g^* = 157.77$ Euro per year and $\tau^* = 0.33$ year. Notice that $g(\tau)$ is not convex, but it is uni-modal. ◇

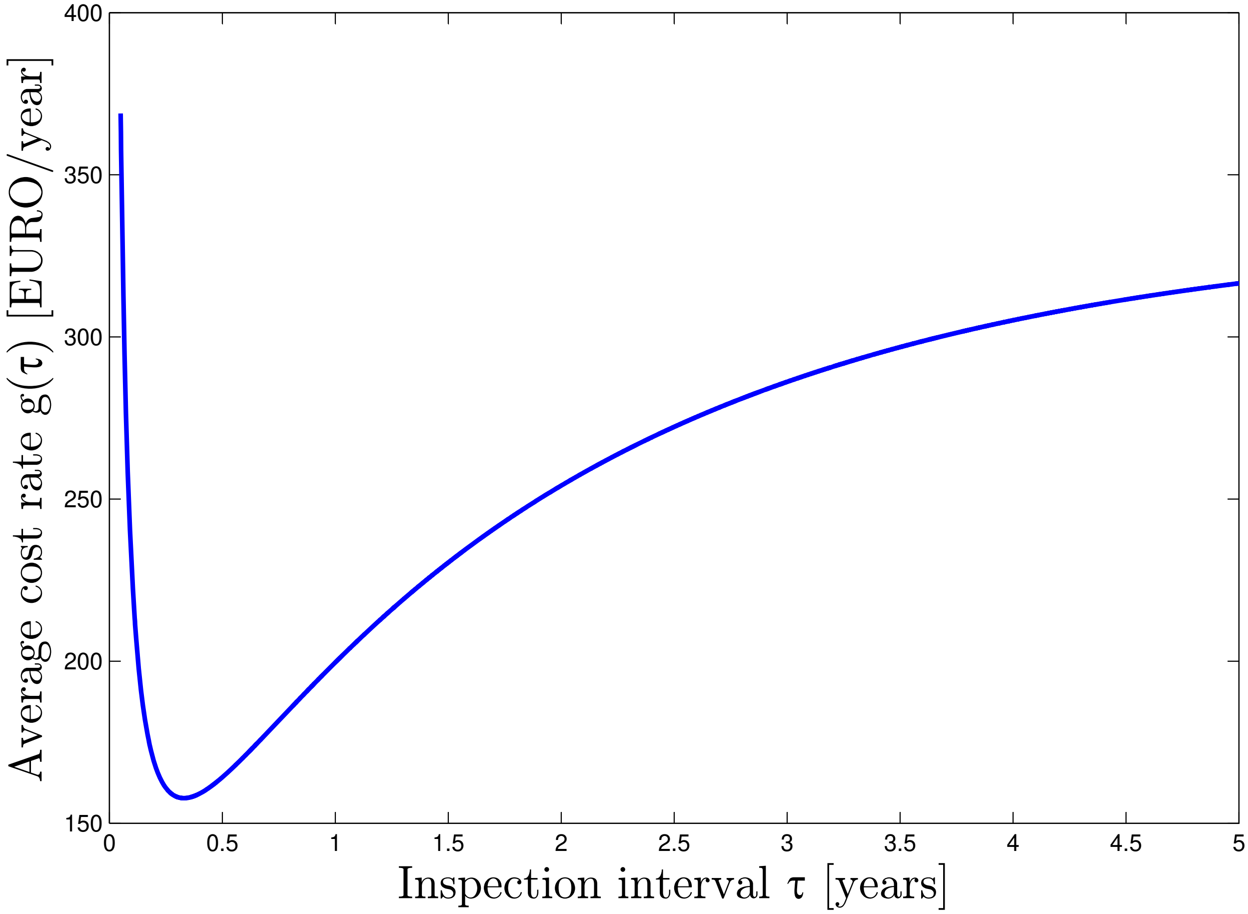


Figure 5.2: Annual cost versus inspection interval for the delay time model in Example 5.1

5.1.2 General time to defect and emergency corrective maintenance

We consider a model that is identical to the model in Section 5.1.1 except that X now has a general distribution rather than an exponential distribution. Refer to Figure 5.1 for the policy we consider. The analysis for this system is quite different, because inspection epochs need no longer be renewal points. (Make sure you see why!) However, each time the component is replaced is a renewal point. Thus there are two type of renewal events: (1) the component is replaced preventively after an inspection, and (2) the component is replaced correctively after failure. The probability that the component is replaced preventively at inspection epoch i since the previous replacement is given by

$$\begin{aligned}
&\mathbb{P}(X+Y > i\tau \cap (i-1)\tau < X < i\tau) \\
&\quad = \int_{x=(i-1)\tau}^{i\tau} \mathbb{P}(X+Y > i\tau \cap (i-1)\tau < X < i\tau \mid X = x) f_X(x)dx \\
&\quad = \int_{x=(i-1)\tau}^{i\tau} \mathbb{P}(Y > i\tau - x) f_X(x)dx \\
&\quad = \int_{x=(i-1)\tau}^{i\tau} (1 - F_Y(i\tau - x)) f_X(x)dx. \qquad (5.10)
\end{aligned}$$

If we let Z denote the time until the next renewal point starting from a renewal point, then its distribution $\mathbb{P}(Z \leq t)$ for $t \in ((i-1)\tau, i\tau)$ and $i \in \mathbb{N}$ is given by

$$\begin{aligned}
\mathbb{P}(Z \leq t) &= \mathbb{P}(X+Y \leq t \cap (i-1)\tau < X < i\tau) + \mathbb{P}(X < (i-1)\tau) \\
&= \int_{x=(i-1)\tau}^{i\tau} \mathbb{P}(X+Y \leq t \cap (i-1)\tau < X < i\tau \mid X = x) f_X(x)dx + F_X((i-1)\tau) \\
&= \int_{x=(i-1)\tau}^{i\tau} \mathbb{P}(Y \leq t - x) f_X(x)dx + F_X((i-1)\tau) \\
&= \int_{x=(i-1)\tau}^{i\tau} F_Y(t-x) f_X(x)dx + F_X((i-1)\tau), \quad i \in \mathbb{N}, \quad t \in ((i-1)\tau, i\tau). \qquad (5.11)
\end{aligned}$$

Now from this we have for the density of Z for $t \in ((i-1)\tau, i\tau)$ and $i \in \mathbb{N}$

$$\begin{aligned}
f_Z(t) &= \frac{d}{dt}\mathbb{P}(Z \leq t) \\
&= \frac{d}{dt}\int_{x=(i-1)\tau}^{i\tau} F_Y(t-x) f_X(x)dx + \frac{d}{dt}F_X((i-1)\tau) \\
&= \int_{x=(i-1)\tau}^{i\tau} f_Y(t-x) f_X(x)dx, \quad t \in ((i-1)\tau, i\tau), \quad i \in \mathbb{N}, \quad t \in ((i-1)\tau, i\tau), \qquad (5.12)
\end{aligned}$$

where the last equality holds by applying Leibniz's rule and noting that $\frac{d}{dt}F_X((i-1)\tau) = 0$.

Combining the results above, we find for the expected cycle length:

$$ECL = \sum_{i=1}^{\infty} \left[\int_{t=(i-1)\tau}^{i\tau} \int_{x=(i-1)\tau}^{i\tau} t f_X(x) f_Y(t-x) dx dt + i\tau \int_{x=(i-1)\tau}^{i\tau} (1 - F_Y(i\tau - x)) f_X(x) dx \right]. \qquad (5.13)$$

Now let us analyze the expected cycle cost. If the component fails in the inspection interval $((i-1)\tau, i\tau)$ for some $i \in \mathbb{N}$, then the incurred costs over the cycle are $(i-1)C_i + C_u$. The probability of failing in the interval $((i-1)\tau, i\tau)$ follow by a derivation similar to (5.11) (verify this yourself):

$$\mathbb{P}(X+Y < i\tau \cap (i-1)\tau < X < i\tau) = \int_{x=(i-1)\tau}^{i\tau} F_Y(i\tau - x) f_X(x)dx. \qquad (5.14)$$

Therefore the expected cost during a cycle due to failure renewal is

$$\sum_{i=1}^{\infty}[(i-1)C_i + C_u]\int_{x=(i-1)\tau}^{i\tau} F_Y(i\tau - x)f_X(x)dx. \tag{5.15}$$

The cost associated with preventive renewal at inspection epoch $i \in N$ is $iC_i + C_p$ so that the expected cost due to preventive renewal is

$$\sum_{i=1}^{\infty}(iC_i + C_p)\int_{x=(i-1)\tau}^{i\tau}(1 - F_Y(i\tau - x))f_X(x)dx. \tag{5.16}$$

Combining the last two equations we find

$$ECC = \sum_{i=1}^{\infty}\Bigg[(iC_i + C_p)\int_{x=(i-1)\tau}^{i\tau}(1 - F_Y(i\tau - x))f_X(x)dx + \\ ([i-1]C_i + C_u)\int_{x=(i-1)\tau}^{i\tau} F_Y(i\tau - x)f_X(x)dx\Bigg]. \tag{5.17}$$

The infinite summation in (5.17) should be truncated at a large number for practical computations. The cost-rate is $g(\tau) = ECC/ECL$. Here too, $g(\tau)$ is not convex, but it usually is uni-modal so that greedy search will yield an optimal inspection interval τ^*.

The performance for a given inspection interval τ for example 5.1 can be computed also using the results in the present section. We leave this up to the reader.

5.1.3 Periodic condition based maintenance

Under some circumstances performing a full maintenance/replacement at any time other than the planned inspection epochs is not feasible. Consider for example a wind-turbine park at sea. To replace/maintain a rotor shaft of a wind-turbine requires the use of equipment such as a jack-up barge; see Figure 5.3. Getting a jack up barge to the wind-turbine park and setting it up is so costly (on the order of 100,000 Euro per day) that it is common practice to maintain multiple wind-turbines at the same time using one jack-up barge. If a wind-turbine fails between inspection epochs, a minimal repair may be applied, for example by sending a small crew on a crew transportation vessel; see Figure 5.4.

We will analyze the following policy for these types of situations: The component is inspected every τ time units. If the component is found to be defect, we apply a preventive replacement at cost C_p. If the component fails during the inspection interval, we apply minimal repair at a cost of C_{mr}. Such a minimal repair restores a component to the condition that is statistically identical to its condition just prior to failure. The last statement is best understood by an example: Suppose

Figure 5.3: A jack up barge is used to replace or conduct maintenance on the rotor shaft or fan blades of a wind-turbine at sea.

Figure 5.4: A crew transportation vessel returning mechanics to shore after the repair of a wind-turbine at sea

$\tau = 4.0$ and a defect occurs at time 1.7 and a failure at time 3.1. After minimal repair, the component is in the same shape as a component that has been operating 3.1-1.7=1.4 time units since a defect occurred. If minimal repair occurs during an inspection interval, then there is a full replacement at the next inspection epoch at the cost C_u. Usually $C_u \geq C_p$ because it is more difficult to maintain a component that has been "kept together" by minimal repairs. (Note that the cost structure is slightly different from what is assumed in Section 3.2.4 for a block policy

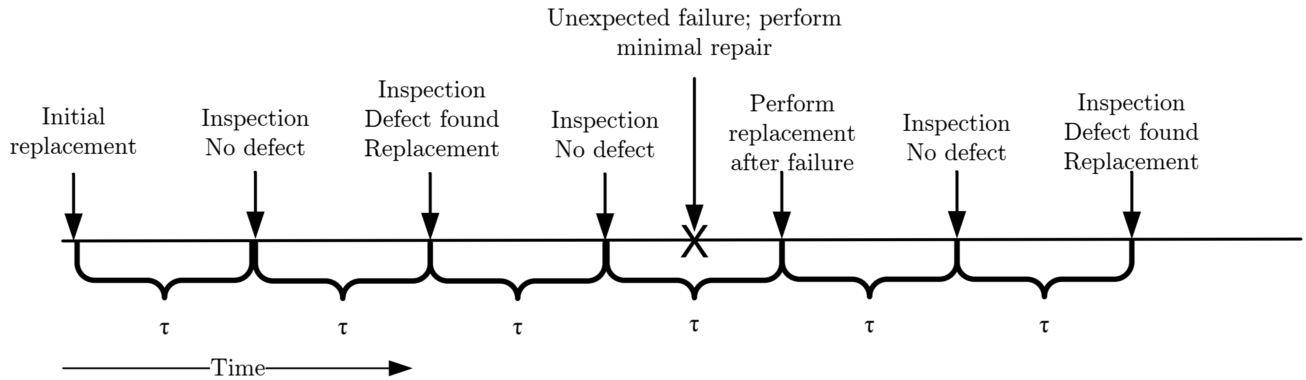


Figure 5.5: Sample path of a policy with minimal repair between inspection epochs and delay time degradation

with minimal repair. That model can easily be modified to the present cost structure if needed.) Finally, there is an inspection cost C_i per inspection epoch. A sample path of this policy is shown in Figure 5.5. We will assume that X, the time to defect, is exponentially distributed. Each inspection epoch is a renewal point under this assumption and the expected cycle length is

$$ECL = \tau. \tag{5.18}$$

The expected cycle costs more involved to determine due to the minimal repairs. If a minimal repair occurs x time units after the last inspection epoch we know that the expected number of minimal repairs is $\int_0^{\tau-x} h_Y(t)dt$. Therefore we can condition on the time of the defect to find that the expected number of minimal repair in a cycle is given by

$$\int_{x=0}^{\tau} \int_{t=0}^{\tau-x} h_Y(t)dt f_X(x)dx.$$

Now it is straightforward to determine the expected cycle costs,

$$\begin{aligned} ECC &= C_{mr} \int_{x=0}^{\tau} \int_{t=0}^{\tau-x} h_Y(t)dt f_X(x)dx + C_u \mathbb{P}(T \leq \tau) + C_p \mathbb{P}(X < \tau \cap X + Y > \tau) + C_i \\ &= C_{mr} \int_{x=0}^{\tau} \int_{t=0}^{\tau-x} h_Y(t)dt f_X(x)dx + C_u F_T(\tau) + C_p \int_{x=0}^{\tau} (1 - F_Y(\tau - x)) f_X(x)dx + C_i, \end{aligned} \tag{5.19}$$

where the last equality follows from using (5.4). The cost-rate of this policy is $g(\tau) = ECC/ECL$ and can be optimized numerically; see the example below.

Example 5.2. Consider a component where X is exponential with mean λ_X^{-1} year, Y is exponential with mean λ_Y^{-1} year, $C_u = 175$, $C_p = 100$, $C_i = 5$, and $C_{mr} = 85$ ($\lambda_X \neq \lambda_Y$). From Example 5.1 we have that

$$\mathbb{P}(X + Y \leq t) = \mathbb{P}(T \leq T) = F_T(t) = 1 - \frac{\lambda_Y e^{-\lambda_X t} - \lambda_X e^{-\lambda_Y t}}{\lambda_Y - \lambda_X},$$

and

$$\mathbb{P}(X \leq \tau \cap X + Y > \tau) = \int_0^\tau (1 - F_Y(\tau - t)) f_X(t) dt = \int_0^\tau \lambda_X e^{-\lambda_X t} e^{-\lambda_Y(\tau - t)} dt = \lambda_X \frac{e^{-\lambda_X \tau} - e^{-\lambda_Y \tau}}{\lambda_Y - \lambda_X}.$$

Next we compute the expected number of minimal repairs in a renewal cycle:

$$\begin{aligned}
\int_{x=0}^{\tau} \int_{t=0}^{\tau - x} h_Y(t) f_X(x) dt dx &= \int_{x=0}^{\tau} \lambda_X e^{-\lambda_X x} \int_{t=0}^{\tau - x} \lambda_Y dt dx \\
&= \int_{x=0}^{\tau} \lambda_Y (\tau - x) \lambda_X e^{-\lambda_X x} dx \\
&= \lambda_Y \tau \mathbb{P}(X \leq \tau) - \lambda_X \lambda_Y \int_0^\tau x e^{-\lambda_X x} dx \\
&= \lambda_Y \tau (1 - e^{-\lambda_X \tau}) - \lambda_Y \frac{1 - e^{-\lambda_X \tau}(\lambda_X \tau + 1)}{\lambda_X} \qquad (5.20)
\end{aligned}$$

The expected cycle costs are now given by

$$\begin{aligned}
ECC = 85 \left[\lambda_Y \tau (1 - e^{-\lambda_X \tau}) - \lambda_Y \frac{1 - e^{-\lambda_X \tau}(\lambda_X \tau + 1)}{\lambda_X} \right] + 175 \left(1 - \frac{\lambda_Y e^{-\lambda_X \tau} - \lambda_X e^{-\lambda_Y \tau}}{\lambda_Y - \lambda_X} \right) \\
+ 100 \lambda_X \frac{e^{-\lambda_X \tau} - e^{-\lambda_Y \tau}}{\lambda_Y - \lambda_X} + 5. \qquad (5.21)
\end{aligned}$$

The expected costs per time unit are shown in Figure 5.6 for $\lambda_X = 0.5$ and $\lambda_Y = 4$. Note that the expected costs per time unit are not convex in τ but they are uni-modal. However, they need not be uni-modal. If $\lambda_Y = 0.1$ for the example above, the optimal inspection interval τ^* approaches infinity. The reason for this is that the delay time is much longer than the time to defect and is memoryless. This is also a general principle: If the delay time of much longer than the time to defect, then condition based maintenance based on defect inspections is not beneficial. The optimal inspection interval for the current example is $\tau^* = 0.22$ and year with an annual cost of 100.19 EURO per year, as can be seen in Figure 5.6. ◇

5.2. Preventive maintenance thresholds for Markovian degradation

We consider a component subject to degradation according to a discrete state Markovian degradation process $X(t)$ that is inspected and available for preventive maintenance at times $0, \tau, 2\tau, 3\tau, \ldots$.

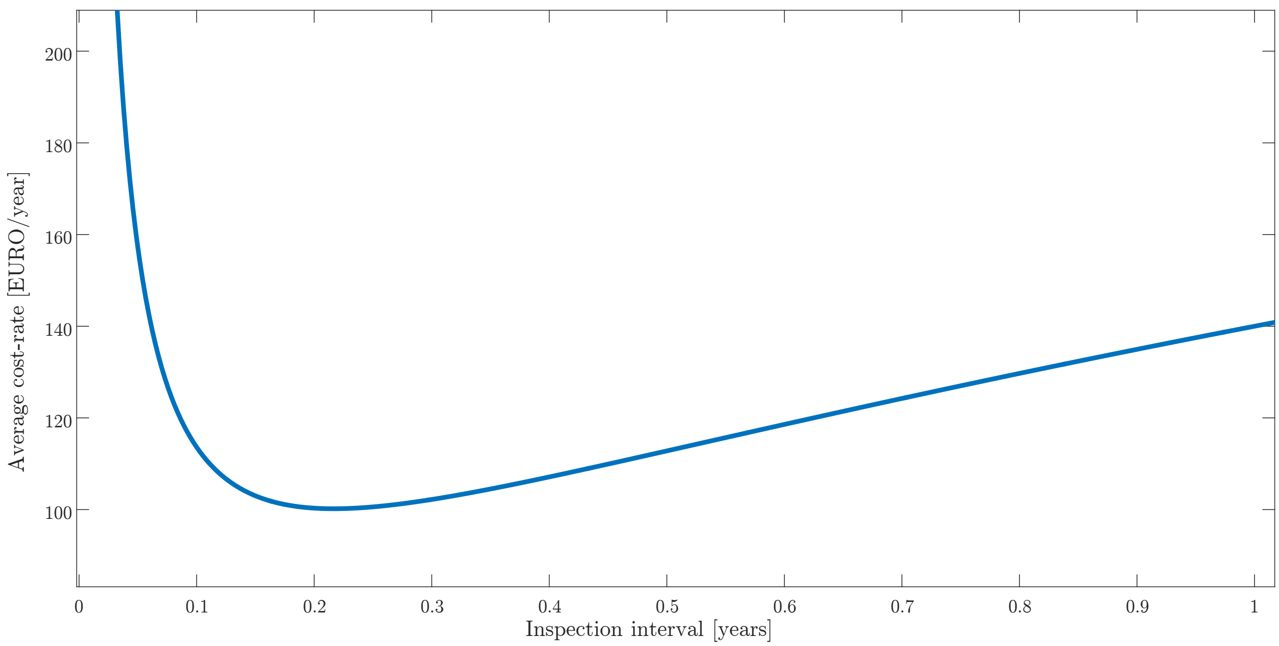


Figure 5.6: Annual cost versus inspection interval for the periodic condition based maintenance model in Example 5.2.

The times $0, \tau, 2\tau, \ldots$ are also called inspection epochs. We assume now that τ is given. With a slight abuse of notation, we will use X_t to denote $X(t\tau)$ for $t \in \mathbb{N}_0$, i.e., X_t denotes the degradation level at inspection epoch $t \in N_0$. The (finite) set of degradation levels is $\mathcal{S} = \{0, 1, 2, 3, \ldots, L\}$, where L denotes the failed state. We let $p_{ij} = \mathbb{P}(X_{t+1} = j \mid X_t = i)$ denote the transition probabilities of X_t given that we do not interfere with the system. If possible, it is customary to order states in increasing order of degradation. For the degradation processes discussed in Chapter 4 this is possible and natural[1]. At each inspection epoch we observe X_t and decide whether we apply preventive replacement/maintenance at cost C_p or continue without performing maintenance. If we find that $X_t = L$, we replace/maintain correctively at cost C_u irrespective of when exactly the component failed between inspection epochs. If we replace/maintain the component, the degradation levels moves to 0 instantaneously and the component will start degrading from that level again. Since the cost of corrective maintenance usually includes the cost of down-time, C_u should be tied to the length of the inspection interval τ (Why?). We will explore this further in Section

[1]Ordering states in a degradation process in some kind of increasing level of degradation is not always possible and it is also not essential as long as there is a failed state L. An example where this is not possible occurs when degradation can follow two different degradation paths in the same Markov process depending on the nature of the initial defect. Within a single path it is possible to order stated in increasing levels of degradation, but making such an ordering between paths is not always possible.

5.2.3 and Chapter 6. We assume that there is no inspection cost C_i because for fixed τ, such cost would be sunk for the decision at hand. Our present goal is to decide at what degradation levels we should replace/maintain preventively at a given inspection epoch so as to minimize the average cost-rate.

When it is possible to order degradation states in increasing order of degradation, it is plausible that there is an optimal *control limit* replacement rule. Under such a rule, there is control limit M, such that a component is replaced at inspection epoch t if $X_t \geq M$. The next subsections give two different algorithms to compute optimal replacement policies, and as a byproduct the optimal control limit M^* if it exists. The existence of such an optimal control limit replacement policy is guaranteed under mild conditions on the degradation process; see Kolesar (1966). These conditions basically state that the state space can be ordered in increasing levels of degradation and that the degradation process should not be able to ameliorate[2] much. For an exact statement of this result, see Kolesar (1966) Theorem 1. All the degradation processes in Chapter 4 are non-decreasing so for them the optimal preventive replacement policy can be described by a control limit. Figure 5.7 shows a sample path of a component subject to negative binomial degradation and a control limit preventive replacement policy. The next two subsection provide two different algorithms to compute optimal replacement decision rules.

5.2.1 Stochastic dynamic programming solution

Stochastic dynamic programming is a very versatile technique to optimize decisions made in systems that evolve over time in a stochastic (Markovian) manner. We will apply this technique to the problem above.

At each inspection epoch t, we see the state X_t, and can decide to either replace the component or not. We let $\mathcal{A}_x$ denote the set of possible decisions in state $x \in \mathcal{S}$ and A_t the decision that we make in period t. For $x < L$, $\mathcal{A}_x = \{0, 1\}$, where we let 1 denote the decision to replace and 0 denote the decision to do nothing. We must replace the component once it has failed so that $\mathcal{A}_L = \{1\}$. Now we let $p_{ij}^a = \mathbb{P}(X_{t+1} = j | X_t = i \cap A_t = a)$, i.e., p_{ij}^a is the probability of the degradation level jumping from i to j at the next inspection epoch if we make decision a at time t. These probabilities can be derived from the degradation process.

Example 5.3. Consider a component subject to Erlangian degradation with 4 degradation states so that $L = 3$ and the state space is given by $\mathcal{S} = \{0, 1, 2, 3\}$. Suppose that the exponential sojourn time in each state has rate $\lambda = 2.0$ and that the inspection interval has length $\tau = 0.5$. The number of degradation steps that occur between inspection epochs is a Poisson random variable with mean

[2] Ameliorate is the opposite of deteriorate or degrade.

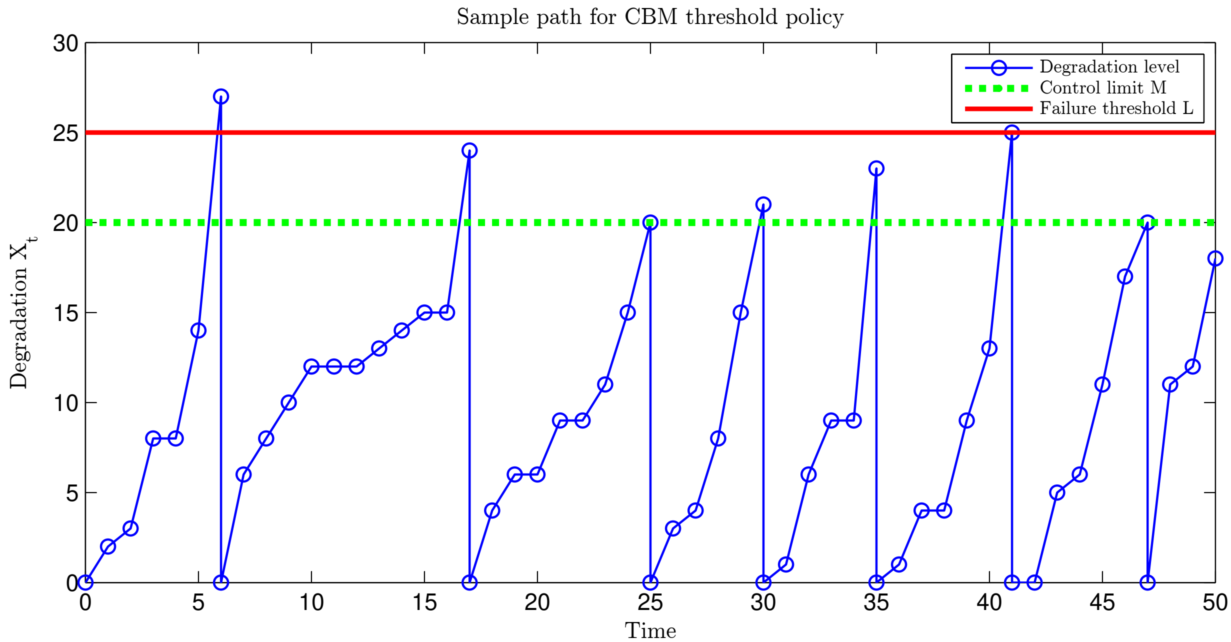


Figure 5.7: Sample path of a system subject to negative binomial degradation that is periodically inspected and replaced/maintained when degradation exceeds the control limit M

$\mu = \lambda\tau = 1$ that is truncated at the failed state $L = 3$. Let Y denote a Poisson random variable with mean $\mu = 1$. Then we have for $p^0_{0,y}$:

$$\begin{aligned}
p^0_{00} &= \mathbb{P}(Y=0) = \frac{(\lambda\tau)^0}{0!}e^{-\lambda\tau} = e^{-2.0\cdot 0.5} = e^{-1} = 0.3679\\
p^0_{01} &= \mathbb{P}(Y=1) = \ldots = 0.3679\\
p^0_{02} &= \mathbb{P}(Y=2) = \ldots = 0.1839\\
p^0_{0L} &= \mathbb{P}(Y \geq 3) = 1 - \mathbb{P}(Y \leq 2) = \ldots = 0.0803. \qquad (5.22)
\end{aligned}$$

Continuing in similar fashion we find that

$$P^0 = \begin{pmatrix} p^0_{00} & p^0_{01} & p^0_{02} & p^0_{0L} \\ p^0_{10} & p^0_{11} & p^0_{12} & p^0_{1L} \\ p^0_{20} & p^0_{21} & p^0_{22} & p^0_{2L} \\ p^0_{L0} & p^0_{L1} & p^0_{L2} & p^0_{LL} \end{pmatrix} = \begin{pmatrix} 0.3679 & 0.3679 & 0.1839 & 0.0803 \\ 0 & 0.3679 & 0.3679 & 0.2642 \\ 0 & 0 & 0.3679 & 0.6321 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad (5.23)$$

For the decision to replace, the transition probabilities are independent of the initial state because replacement means instantaneously moving to state 0:

$$P^1 = \begin{pmatrix} p^1_{00} & p^1_{01} & p^1_{02} & p^1_{0L} \\ p^1_{10} & p^1_{11} & p^1_{12} & p^1_{1L} \\ p^1_{20} & p^1_{21} & p^1_{22} & p^1_{2L} \\ p^1_{L0} & p^1_{L1} & p^1_{L2} & p^1_{LL} \end{pmatrix} = \begin{pmatrix} 0.3679 & 0.3679 & 0.1839 & 0.0803 \\ 0.3679 & 0.3679 & 0.1839 & 0.0803 \\ 0.3679 & 0.3679 & 0.1839 & 0.0803 \\ 0.3679 & 0.3679 & 0.1839 & 0.0803 \end{pmatrix}. \tag{5.24}$$

Observe that these transition probabilities can easily be computed for different inspection intervals. ◇

Now let us embark on the following thought experiment: Suppose that we only need to make decisions for a finite number of inspection epochs, N say. After inspection epoch N, the system just stops and no more decisions need to be made. However, if the component is found to have failed at inspection epoch N, i.e., $X_N = L$, we still need to pay C_u for the corrective replacement / maintenance. Now we introduce some crucial notation: Let $V_n(x)$ denote the minimal expected cost that will be incurred from period $N-n$ up to and including period N when $X_{N-n} = x$. $V_n = (V_n(0), V_n(1), V_n(2), \ldots, V_n(L))^T$ is also referred to as the value function because $V_n(x)$ expresses the expected value (cost) of being in state x with n periods to go. Now by definition $V_0(x) = 0$ for $x < L$ and $V_0(L) = C_u$. The vector $(V_0(0), V_0(1), \ldots, V_0(L))^T$ is also referred to as the terminal cost vector or terminal value function. Suppose now that we are in period $N-1$ in state $X_{N-1} = x$ and need to decide what to do. If $x = L$, we only have one possible decision (replace) and so

$$V_1(L) = C_u + p^1_{L,L} C_u. \tag{5.25}$$

If $x < L$, we can either replace or not replace leading to the following two expected costs:

$$\text{replace: } C_p + p^1_{x,L} C_u, \qquad \text{do not replace: } p^0_{x,L} C_u. \tag{5.26}$$

Since $V_1(x)$ is the minimal expected cost, it is the minimum of these two.

$$V_1(x) = \min(C_p + p^1_{x,L} C_u, p^0_{x,L} C_u). \tag{5.27}$$

Recalling the values of the terminal cost vector, (5.25) and (5.27) can be combined to

$$V_1(x) = \begin{cases} \min\{C_p + \sum_{y \in \mathcal{S}} p^1_{x,y} V_0(y), \quad \sum_{y \in \mathcal{S}} p^0_{x,y} V_0(y)\} & x < L; \\ C_u + \sum_{y \in \mathcal{S}} p^1_{x,y} V_0(y), & x = L. \end{cases} \tag{5.28}$$

Now suppose that we are in period $N-2$ in state $X_{N-2} = x$ and we need to make the decision whether or not to replace and we will make optimal decisions in period $N-1$. Then the decision to replace leads to a total cost from period $N-2$ up to N of

$$C_p + \sum_{y \in \mathcal{S}} p^1_{xy} V_1(y), \tag{5.29}$$

whereas the decision to not replace leads to a cost of

$$\sum_{y\in\mathcal{S}} p^0_{xy} V_1(y). \tag{5.30}$$

Since we must replace in state L and $V_2(x)$ is the optimal cost to go with 2 periods left to the end of the horizon, we have

$$V_2(x) = \begin{cases} \min\{C_p + \sum_{y\in\mathcal{S}} p^1_{x,y} V_1(y), \quad \sum_{y\in\mathcal{S}} p^0_{x,y} V_1(y)\} & x < L; \\ C_u + \sum_{y\in\mathcal{S}} p^1_{x,y} V_1(y), & x = L. \end{cases} \tag{5.31}$$

Now this argument can be repeated for $N-3$, $N-4$ and so forth to find the general expression

$$V_n(x) = \begin{cases} \min\{C_p + \sum_{y\in\mathcal{S}} p^1_{x,y} V_{n-1}(y), \quad \sum_{y\in\mathcal{S}} p^0_{x,y} V_{n-1}(y)\} & x < L; \\ C_u + \sum_{y\in\mathcal{S}} p^1_{x,y} V_{n-1}(y), & x = L. \end{cases} \tag{5.32}$$

Using (5.32), we can compute the optimal expected cost with n inspection epochs to go until we stop operating the system. Equation (5.32) is known as the *Bellman optimality equation* and also as the *Dynamic programming recursion.* From the value function V_n, we can deduce the optimal replacement decision with n periods to go in state $x < L$: It is optimal to replace if

$$C_p + \sum_{y\in\mathcal{S}} p^1_{x,y} V_{n-1}(y) \le \sum_{y\in\mathcal{S}} p^0_{x,y} V_{n-1}(y) \tag{5.33}$$

and to not replace if the inequality above is reversed. This also means that if there is an optimal threshold with n periods to go, say M^*_n, it satisfies

$$M^*_n = \min\left\{ x \in \mathcal{S} \mid C_p + \sum_{y\in\mathcal{S}} p^1_{x,y} V_{n-1}(y) \le \sum_{y\in\mathcal{S}} p^0_{x,y} V_{n-1}(y) \right\}. \tag{5.34}$$

By Theorem 1 of Kolesar (1966), such a control limit M^*_n always exists if the degradation process is non-decreasing. All the degradation processes in Chapter 4 have this property.

Example 5.4. Reconsider Example 5.3 and suppose that $C_u = 1000$ and $C_p = 300$ Euro. Using the probabilities computed in Example 5.3, we can compute $V_n(x)$ for $x \in \mathcal{S}$ and $n = 0, 1, \ldots$. The results for $n = 0, \ldots, 10$ are shown in Table 5.1 and Figure 5.8. From these we may compute M^*_n for $n = 0, \ldots, 10$. This result is shown in Figure 5.9. It appears that a control limit of 2 is optimal for all inspection epochs preceding sudden termination of the system. Furthermore notice from Figure 5.8 that the distance between $V_n(x)$ and $V_{n-1}(x)$ appears to converge to the same number for each $x \in \mathcal{S}$. ◇

Table 5.1: Value function $V_n(x)$ for Example 5.4

x	$V_0(x)$	$V_1(x)$	$V_2(x)$	$V_3(x)$	$V_4(x)$	$V_5(x)$	$V_6(x)$	$V_7(x)$	$V_8(x)$	$V_9(x)$	$V_{10}(x)$
0	0.00	80.30	283.45	506.90	730.35	953.81	1177.26	1400.71	1624.16	1847.61	2071.06
1	0.00	264.24	522.57	746.03	969.48	1192.93	1416.38	1639.83	1863.28	2086.73	2310.18
2	0.00	380.30	583.45	806.90	1030.35	1253.81	1477.26	1700.71	1924.16	2147.61	2371.06
L	1000.00	1080.30	1283.45	1506.90	1730.35	1953.81	2177.26	2400.71	2624.16	2847.61	3071.06

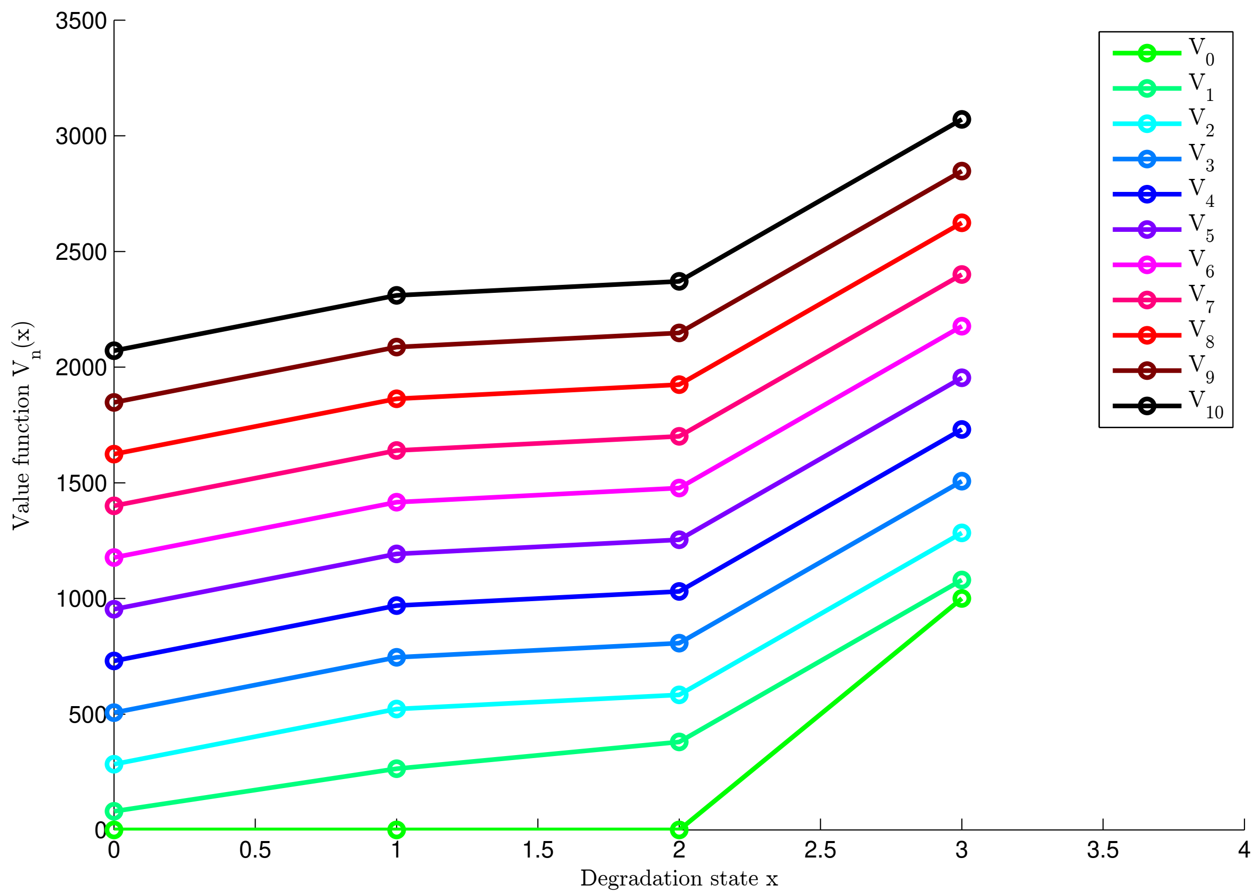


Figure 5.8: Value functions V_0 to V_{10} for Example 5.4.

Example 5.5. Reconsider Examples 5.3 and 5.4. Suppose now that the Erlang degradation process has 41 phases so that $L = 40$. What is the behaviour of M_n^* as a function of n for $\lambda = 2, 5, 10, 15$? This behaviour is shown in Figure 5.10. It appears that the thresholds converge to a single threshold as the number of periods to go increases.

We are not looking for decisions that minimize the cost with n inspection epochs to go, but

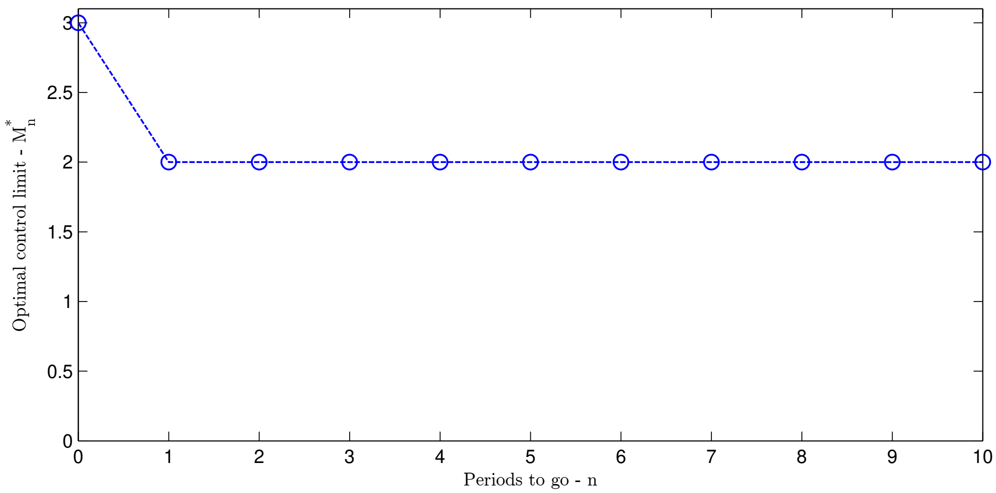


Figure 5.9: Optimal control limits for Example 5.4 with n periods to go

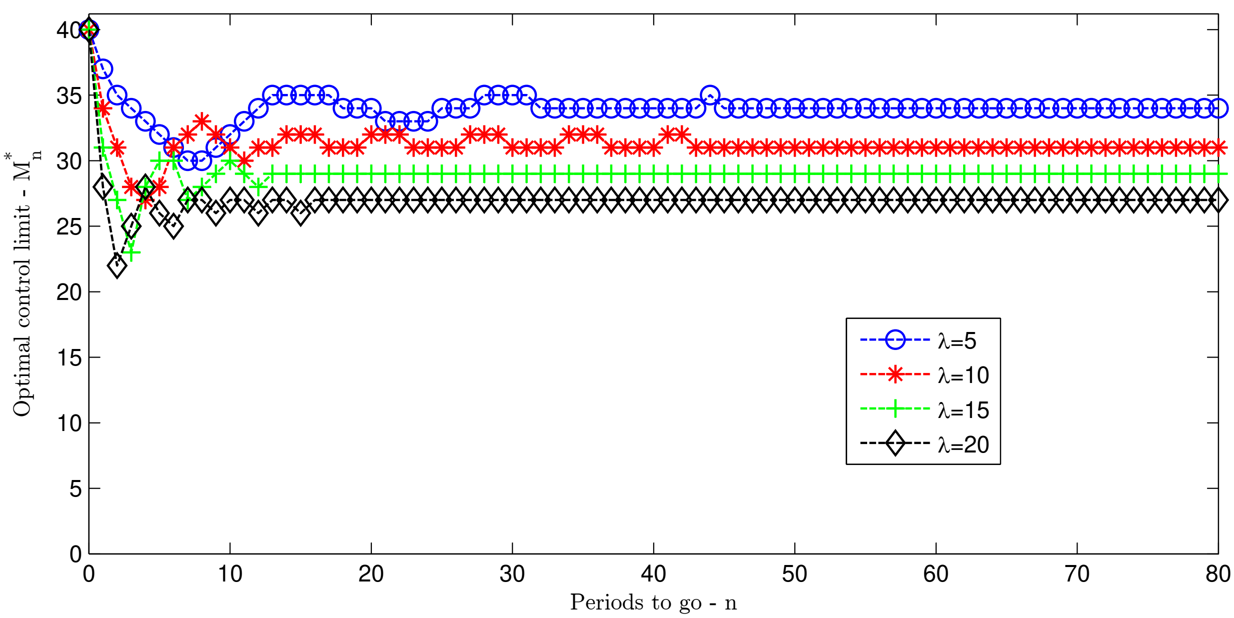


Figure 5.10: Optimal control limits as n grow for different λ

for decisions that minimize the average cost-rate. It seems reasonable to expect that the decisions that are optimal with n periods to go, for some very large n, are decisions that are good on average. This is what the results in Example 5.5 suggest at least. Under some mild conditions that all the degradation processes in Chapter 4 satisfy[3], the optimal replacement decisions with n periods to go converge to decisions that minimize the average cost-rate as n approaches infinity. So we can

[3]The mild conditions are satisfied if a certain degradation state is reachable for any conceivable maintenance

use Bellman equation (5.32) to find replacement policies that are optimal for the average cost-rate. But how do we know whether the n we choose is large enough, and what is the average cost-rate that these decisions achieve? To say something about this we need some more notation. Let

$$g_n = \frac{\max_{x\in\mathcal{S}}(V_n(x) - V_{n-1}(x)) + \min_{x\in\mathcal{S}}(V_n(x) - V_{n-1}(x))}{2}. \tag{5.35}$$

Thus g_n represents the average of the maximum and minimum cost increase from $n-1$ inspection epochs to go, to n inspection epochs to go. It can be interpreted as an estimate of the distance between V_n and V_{n-1}, which is easy to visualize looking at Figure 5.8. Let g'_n be the average cost per inspection interval of applying the replacement decisions corresponding to control limit M^*_n and let g^* be the cost-rate per inspection interval of a control limit policy that minimizes the average cost-rate per inspection epoch. From the theory of stochastic dynamic programming (also known as Markov decision processes), we have the following theorem (Puterman, 1994, Theorem 8.5.4 and 8.5.6):

Theorem 5.1. *For any $\varepsilon > 0$ there is an $n < \infty$ such that*

$$\max_{x\in\mathcal{S}}(V_n(x) - V_{n-1}(x)) - \min_{x\in\mathcal{S}}(V_n(x) - V_{n-1}(x)) < \varepsilon. \tag{5.36}$$

For this n, we have $|g_n - g'_n| < \varepsilon/2$, $|g'_n - g^*| < \varepsilon/2$, and $|g_n - g^*| < \varepsilon$. Thus the average cost-rate of applying control limit M^*_n can exceed the optimal cost-rate by no more than ε and g_n is within ε of the optimal cost-rate.

Based on this theorem, we can find a control limit that is as good as we like by the following algorithm:

1. Set $\varepsilon > 0$, $n = 1$, $V_0(L) = C_u$ and $V_0(x) = 0$ for all $x < L$.

2. Compute V_n using (5.32). If (5.36) is satisfied, continue to the next step, otherwise set $n = n + 1$ and repeat this step.

3. Compute M^*_n and g_n using (5.34) and (5.35) and output these as the optimal control limit and cost per inspection interval respectively.

Note that the average cost per inspection interval can be translated to the expected cost per time unit by dividing by τ. Common practice is to set $\varepsilon = 10^{-6}$.

policy and this state has transitions to itself. Under all degradation processes we consider in Chapter 4, L and 0 are reachable under any policy and there are transitions from 0 to itself so these conditions are verified.

Example 5.6. Reconsider Examples 5.3, 5.4, and 5.5. Using the algorithm above, we compute policies that are within $\varepsilon = 10^{-6}$ of optimality. The results are shown in Table 5.2 below for all the instances considered in Examples 5.4 and 5.5, and one additional instance of a component subject to Erlangian degradation. We see that the optimal threshold is indeed given by the thresholds to which we see convergence in Figure 5.10. Now we also know the average cost-rates that these control limits induce. $\diamond$

Table 5.2: Optimal control limits and costs for several components subject to Erlangian degradation.

Instance	$\lvert\mathcal{S}\rvert$	L	λ	τ	C_u	C_p	M^*	g^*
1	4	3	2	0.5	1000	300	2	223.45
2	41	40	5	0.5	1000	300	34	21.67
3	41	40	10	0.5	1000	300	31	45.90
4	41	40	15	0.5	1000	300	29	72.16
5	41	40	20	0.5	1000	300	27	100.71
6	51	50	3	3	5000	900	36	208.51

5.2.2 Linear programming solution

Optimal control limits for the problem in Section 5.2 can also be found by linear programming. We will use the notation introduced in the first paragraph of Section 5.2.1 and use the word *policy* to refer to a decision rule. A replacement policy can be encoded as follows: Let $f_{x,a} = \mathbb{P}(A_t = a \mid X_t = x)$ denote the probability of applying action $a \in \mathcal{A}(x)$ when the degradation level is x during an inspection epoch. If $f_{x,1} = 1$, we always replace the component when finding degradation x during an inspection epoch, if $f_{x,0} = 1$, we never replace the component when finding degradation x during an inspection epoch, and if $f_{x,1} = u \in (0,1)$, then we replace with probability u and do not replace with probability $f_{x,0} = 1 - u$ when finding degradation x during an inspection epoch. Next let $\pi_x = \lim_{t\to\infty} \mathbb{P}(X_t = x)$ denote the steady state probability of finding degradation level x during an inspection epoch when the system is controlled by the replacement policy encoded by $f_{x,a}$. Then $\pi_x f_{x,a}$ is the steady state probability finding degradation level x during an inspection and making decision a. Let us denote this probability

$$z_{x,a} = \lim_{t\to\infty} \mathbb{P}(X_t = x \cap A_t = a) = \pi_x f_{x,a}. \tag{5.37}$$

For example, $z_{5,0}$ is the steady state probability of finding a degradation level of 5 during an inspection epoch and deciding not to replace the component. Observe that

$$\pi_x = \sum_{a \in \mathcal{A}_x} z_{x,a}. \tag{5.38}$$

Under our notation, the average cost per inspection interval can be written as

$$g = \sum_{x \in \mathcal{S} \setminus \{L\}} C_p z_{x,1} + C_u z_{L,1}. \tag{5.39}$$

The balance equations for the Markov process X_t under this policy are

$$\pi_x = \sum_{y \in \mathcal{S}} \sum_{a \in \mathcal{A}_y} \pi_y f_{y,a} p^a_{y,x}, \quad x \in \mathcal{S}, \tag{5.40}$$

and using (5.37) and (5.38) these equations can be rewritten as

$$\sum_{a \in \mathcal{A}_x} z_{x,a} = \sum_{y \in \mathcal{S}} \sum_{a \in \mathcal{A}_y} z_{y,a} p^a_{y,x}, \quad x \in \mathcal{S}. \tag{5.41}$$

Now any set of $z_{x,a}$ that satisfy (5.41) and normality ($\sum_{x \in \mathcal{S}} \sum_{a \in \mathcal{A}_x} z_{x,a} = 1$) corresponds to a policy where the steady state probabilities and decisions can be recovered by using (5.37) and (5.38):

$$\pi_x = \sum_{a \in \mathcal{A}_x} z_{x,a}, \qquad f_{x,a} = \frac{z_{x,a}}{\sum_{b \in \mathcal{A}_x} z_{x,b}}. \tag{5.42}$$

Therefore the cost of an optimal replacement policy can be found by choosing $z_{x,a}$ to minimize (5.39) while satisfying (5.41) and normality, which is the following linear program:

$$\begin{aligned}
(\mathbf{LP}) \qquad \min_{z_{x,a}} \qquad & g = \sum_{x \in \mathcal{S} \setminus \{L\}} C_p z_{x,1} + C_u z_{L,1} \\
\text{subject to} \qquad & \sum_{a \in \mathcal{A}_x} z_{x,a} = \sum_{y \in \mathcal{S}} \sum_{a \in \mathcal{A}_y} z_{y,a} p^a_{y,x}, \quad \forall x \in \mathcal{S} \\
& \sum_{x \in \mathcal{S}} \sum_{a \in \mathcal{A}_x} z_{x,a} = 1 \\
& z_{x,a} \geq 0, \quad \forall x \in \mathcal{S}, \quad \forall a \in \mathcal{A}_x.
\end{aligned} \tag{5.43}$$

After (5.43) has been solved to optimality, the objective g^* represents the optimal cost per inspection interval and an optimal policy can be obtained from the optimal $z^*_{x,a}$ using (5.42). From the theory of linear programming, it can be shown that for each $x \in \mathcal{S}$, either $z_{x,0} > 0$ or $z_{x,1} > 0$, but not both. Therefore, the optimal control limit M^* satisfies

$$M^* = \min\{x \in \mathcal{S} \mid z^*_{x,1} > 0\}, \tag{5.44}$$

if it exists.

Remark 5.1. Some care must be taken in solving the linear program above. The solution can only be trusted if $\sum_{a\in\mathcal{A}(x)} z^*_{x,a} > 0$ for all $x \in \mathcal{S}$. A sufficient condition for this is the existence of a degradation state that can be reached under any maintenance policy. The degradation processes in Chapter 4 satisfy this condition, because each state in $\mathcal{S}$ can be reached under any policy.

Remark 5.2. Stochastic dynamic programming and linear programming are general approaches for making decisions in a Markov process whose cost and evolution depends on these decision. This field is also called Markov decision processes (MDP) and stochastic dynamic programming and linear programming are common solution algorithms. Other applications of MDP include inventory and supply chain management, manufacturing, and call centers. For example, the state of an inventory problem is often the inventory position and the decisions are ordering decisions.

Example 5.7. Reconsider Examples 5.3 and 5.4 where degradation occurs as an Erlangian process with 4 states, sojourn times in each state are exponential with rate $\lambda = 2$, the inspection interval $\tau = 0.5$, $C_p = 300$, and $C_u = 1000$. We can solve the linear program (5.43) for this instance to find the results in Table 5.3. We can read from the Table that $M^* = 2$, which coincides with what we found in Example 5.6. The optimal objective also coincides: $g^* = 223.45$. Note that since g^* is the optimal cost per inspection epoch, the optimal cost per time unit is $g^*/\tau = 446.90$. $\diamond$

Table 5.3: Solution of linear program (5.43) for instance of Example 5.7

	$z^*_{x,a}$			$f^*_{x,a}$	
x	$a = 0$	$a = 1$	π^*_x	$a = 0$	$a = 1$
0	0.232544	0	0.232544	1	0
1	0.367879	0	0.367879	1	0
2	0	0.251607	0.251607	0	1
3 (L)	-	0.147969	0.147969	-	1

5.2.3 Extensions

The model of Section 5.2 can be extended in several ways. Here we briefly sketch several directions.

The inspection interval need not be given a priori. In this case, the inspection interval must be optimized as well. Some care must be taken here to ensure that C_u should now become an increasing function of τ for this to make any sense. (If C_u remains fixed, then as $\tau \to \infty$, $g \to 0$ and that does not make physical sense. In essence we can let the component break down, but never pay for the replacement because we simply do not inspect the component to find that it has broken.) Inspection costs C_i are no longer sunk when τ is a decision variable, so these need to be included as well. If we assume that component breakdown leads to down-time, and down-time is penalized linearly in the amount of down-time, a good choice is to let $C_u(\tau) = C_1\tau + C_2$ where C_1 captures the down-times cost times the expected amount of time a component was down per time unit in a failure interval, and C_2 captures the actual cost of a corrective replacement. We can now use the following approach: Compute the optimal control limit for a given τ using either stochastic dynamic programming or linear programming, and then vary τ to find the best τ.

The model in Section 5.2 assumes that the current degradation level carries sufficient information to understand future degradation, at least probabilistically. When degradation is non-linear, it often depends also on the time since the last replacement. In this case, the state should be extended to include the number of inspection epochs since the last replacement. The approach still works for this model, but all the computations become more involved. See Benyamini and Yechiali (1999) for more details.

Finally, there are situations where we would like to determine a maintenance threshold and learn form the degradation path of the component that is currently in use, what the actual transition probabilities are. Such models are also studied, see for example Elwany et al. (2011) and Zhang et al. (2015).

5.3. Exercises

Exercise 5.1. *Old exam exercise about inspections*

Consider a machine that has only one critical component. Before the component actually fails, a potential failure symptom occurs that can be detected via inspection. Such a signal can be, e.g. excessive vibration that can be detected with a counter. This means that the machine can be in three different states: the perfect functioning state, a defective state (in which the failure symptom is detectable), and the failure state. The time until a failure symptom occurs is exponentially distributed with a constant rate of $\lambda = 2$/year, and the time between the occurrence of the failure symptom and the actual failure, i.e. the “delay time” $Y = 0.2$ years, is constant.

The management is considering two maintenance policies for this machine. The first policy considered is the failure-based maintenance policy (FB). The costs of a corrective maintenance actions is $C_u = 7000$ Euro.

(a) What is the average cost per year for the FB policy, g_{FB}?

Policy 2 uses the possibility of preventive maintenance (PM), and can be described as follows: An inspection takes place every $\tau \leq Y = 0.2$ years (i.e., the inspection interval is shorter than the delay time), which costs $C_i = 200$ Euro. The inspections are perfect in the sense that if the system is in a defective state, this will be identified by the inspection. If a defect is identified at an inspection, the component is replaced by a new one at a cost of $C_p = 1000$ Euro.

(b) What is the long term expected cost per year of this policy as a function of τ?

(c) What is the optimal inspection interval, τ^*, of this policy?

Finally, suppose that the delay time is zero ($Y = 0$) and $C_i = 200$ still.

(d) What is the optimal policy in this case? Why?

Exercise 5.2. *Computation of transition probabilities*

Consider a component subject to degradation that is inspected periodically. The degradation level takes values in $\mathcal{S} = \{0, 1, 2, \ldots, L\}$ where L denotes the failed state and X_t denotes the degradation level at the t-th inspection epoch. If no maintenance is applied, then the degradation evolves according to

$$X_{t+1} = \min(X_t + Z_t, L)$$

where $(Z_t)_{t=0}^{\infty}$ are i.i.d. random variables on the integers with distribution $p_z = \mathbb{P}(Z_t = z)$ and $p_{\geq z} = \mathbb{P}(Z_t \geq z) = 1 - \sum_{x=0}^{z-1} p_x$. If maintenance is applied, then the degradation evolves according to

$$X_{t+1} = \min(Z_t, L).$$

Let $p^1_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i \cap \text{maintenance is applied at time } t)$ and $p^0_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i \cap \textit{no} \text{ maintenance is applied at time } t)$.

(a) Suppose that degradation follows an Erlang process with rate μ and that the inspection interval has length τ. Give an expression for p_z in this case.

(b) Suppose that degradation follows a compound Poisson process with arrival intensity λ and logarithmic compounding distribution with parameter q. Suppose further that the inspection interval has length τ. Give a closed form expression for p_z for this case.

(c) Give closed form expressions of p^0_{ij} for $i, j \in \mathcal{S}$ in terms of p_z and $p_{\geq z}$. Use case distinctions if convenient.

(d) Give closed form expressions of p^1_{ij} for $i, j \in \mathcal{S}$ in terms of p_z and $p_{\geq z}$. Use case distinctions if convenient.

Exercise 5.3. *Exam level exercise about comparison delay time model and age-based policies.*

Consider the radio communication system on a helicopter. This system has two antennas that can be used both for transmission and receiving. However, only one of the antennas is used at any given time, that is, the second will only be used after the first has failed. The lifetimes of these antennas has an exponential distribution with a mean of 4 years and the lifetime of an antenna does not start until after it is put into use.

(a) If we do not interfere with the antennas, what is the distribution of the time until both antennas fail?

(b) Suppose that an inspection reveals that one of the two antennas has already failed. Does it make sense to replace both antennas preventively? Why (not)?

(c) Suppose inspections of antennas costs 500 Euro and that replacing one antenna costs 3400 Euro. We inspect both antennas every τ time units and if we observe that one has failed we replace it. If both antennas fail before the next planned inspection, the communication system will stop functioning and this will be noticed immediately. In this case, both antennas will be replaced (for 3400 Euro each) and an additional cost of 11500 Euro will be incurred due to unplanned maintenance. The next inspection after such an incident will be planned τ time units later. Give a closed form expression for the expected cost per time unit under this policy as a function of τ.

(d) Determine the optimal inspection interval τ^* and optimal expected cost per time unit. You may use a computer if convenient.

(e) A team of maintenance managers and engineers would like to compare the policy of question (c) with an age-based policy. Under this age-based policy, the antennas are replaced in full if both fail and otherwise one of them can be replaced preventively when they reach an age of τ. Both antennas have exponential lifetimes so when they reach a certain age, the inspection of 500 Euro is used to determine which antenna needs to be replaced for 3400 Euro. If both antennas are still working, then no antenna will be replaced. Three engineers come up with different suggestions on how to compute the cost of the best age-based policy for the antennas. They are:

(1) Compute the cost of the best age-based policy using the model in Section 3.2.2 with the life-time distribution as determined in question (a) with a planned maintenance cost of $C_p = 3400 + 500 = 3900$ Euro and an unplanned maintenance cost of $C_u = 2 \cdot 3400 + 11500 = 18300$ Euro.

(2) Compute the cost of the best age-based policy using the model in Section 3.2.2 with the life-time distribution as determined in question (a) with a planned maintenance cost that is a function of the preventive replacement time parameter τ as follows: $C_p = 500 + \mathbb{P}(X < \tau)3400 = 500 + (1 - e^{-\tau/4})3400$ Euro and unplanned maintenance cost of $C_u = 2 \cdot 3400 + 1150 = 18300$ Euro. (X is the time it takes for the first antenna to fail.)

(3) Compute the cost of the best age-based policy using the model in Section 3.2.2 with the life-time distribution as determined in question (a) with a planned maintenance cost that is a function of the preventive replacement time parameter τ as follows: $C_p = 500 + \mathbb{P}(X < \tau \mid X + Y > \tau)3400$ Euro and unplanned maintenance cost of $C_u = 2 \cdot 3400 + 1150 = 18300$ Euro. (X and Y are the lifetime of the first and second antenna respectively.)

Should you follow the suggestion of engineer 1, 2, or 3? Why?

(f) Do you expect the cost of the optimal age-based policy as selected in (d) to be strictly more expensive, strictly cheaper or exactly as costly as the optimal policy computed under (c)? Why? (You do not have to perform computations here.)

Exercise 5.4. *Use of a computer is advised for this exercise*

Consider an emergency break system subject to Erlangian degradation with 7 degradation states numbered 0,1,. . . ,6. The Sojourn time in each state is exponential with a mean of 20 weeks. Unless the emergency break is used, a failure cannot be detected without an inspection. Law requires that this break system, along with other safety critical systems on the vehicle, are inspected every 12 weeks by an independent inspection agency. They report the degradation level of the emergency break to the owner. If they find that the emergency break has failed, a safety enforcement fee of 15000 Euro is charged. Before returning a vehicle to service, the emergency break can be replaced at a cost of 4200 Euros regardless of whether it is has failed already or not.

(a) Let p^0_{ij} denote the probability of moving from degradation state i to degradation state j between inspection epochs. Compute p^0_{ij} for $i, j = 0, \ldots 6$.

(b) Let p^1_{ij} denote the probability of moving from current degradation state i to degradation state j when replacement is applied.

(c) In order to minimize the average cost-rate, at what degradation level should the emergency break be replaced? Use both value iteration and linear programming to compute the optimal control limit threshold M^* and average cost per inspection epoch g^*. Verify that your implementations are correct by obtaining the same answer using both methods.

Exercise 5.5. *Use of a computer is advised for this exericse*

Consider a component subject to degradation with 11 degradation states numbered 0,...,10. This system has been inspected weekly for several years now and it was found that the degradation increases each week by 1.27 states on average. The standard deviation of the increase is 1.31.

(a) The increments in the degradation per week can be modeled by a negative-binomial distribution. Estimate the shape and scale of this negative binomial distribution.

(b) Suppose that the degradation process is a compound Poisson process with logarithmic compounding distribution. (This leads to the negative-binomial increments assumed in (a).) Estimate the arrival intensity λ of the Poisson process and the parameter q of the logarithmic compounding distribution.

(c) Let p^0_{ij} denote the probability of moving from degradation state i to j in a week when we do not replace the component. Compute p^0_{ij} for $i, j = 0, \ldots, 10$.

(d) Suppose that applying maintenance return the component to state 0. Let p^1_{ij} denote the probability of moving from state i to state j if we apply maintenance.

(e) Suppose now that applying maintenance will return the component to degradation level 0 with probability $\frac{1}{2}$, but will only return it to degradation level 1 with probability $\frac{1}{2}$. Let X denote the increment of degradation in a week and let B denote a random variable that is 0 with probability $\frac{1}{2}$ and 1 with probability $\frac{1}{2}$. Show that now $p^1_{ij} = \mathbb{P}(\min(X + B, 10) = j)$.

(f) Use the result in (e) to compute p^1_{ij}.

(g) Suppose that a preventive replacement costs 1300 Euro and a corrective replacement costs 6100 Euro. Suppose again that replacement leads to a component with an initial degradation level of 0. Starting from which degradation level should preventive replacement be applied? What is the average cost per week under such a replacement policy?

Exercise 5.6. *Formulation of problem. Exam level*

Consider a cutting tool subject to degradation. When the cutting tool is new, it can cut workpieces at a rate of 7000 per week. These workpieces are sold with at a price of 2 EuroS per workpiece. The demand for these pieces is sufficiently big that all workpieces that can be produced can also be sold. However, this tool is subject to degradation: The speed with which it can cut workpieces degrades over time. The best degradation measure is therefor the number of workpieces that the machine can produce per week, which takes values in $\{0, 1, \ldots, 7000\}$. Each weekend, the factory shuts down and it is possible to replace the cutting tool at a cost of 350 Euro. Suppose that the transition probabilities p_{ij}^a are given; p_{ij}^a denote the transition probability of moving from cutting speed i to cutting speed j in a week if we make replacement decision $a \in \{0, 1\}$, where $a = 1$ denotes replacement and $a = 0$ denotes leaving the current cutting tool in place. Further, you may assume that when the cutting tool can cut at speed x at the beginning of the week, it will be able to cut at this speed all week, before its cutting speed goes down for the next week.

(a) Let $R(x, a)$ denote the profit (revenue minus cost) made during a week when the cutting speed at the end of the week before is x and we take decision $a \in \{0, 1\}$ where $a = 1$ denotes replacement and $a = 0$ denotes the decision not the replace the cutting tool. Give a closed form expression for $R(x, a)$.

(b) Suppose we would like to maximize the weekly profit. Give a dynamic programming formulation of this problem, i.e., formulate the dynamic programming recursion. You do not need to perform any computations.

(c) Give a linear programming formulation for the problem described in (b). You do not need to perform any computations.

5.3.1 Solutions

Solution 5.1. (a) The average annual costs g_{FB} are: $g_{FB} = \frac{ECC}{ECL} = \frac{7000}{0.5+0.2} = 10000$ Euro per year.

(b) If $\tau < 0.2$ then a failure can never occur so in this case $ECC = C_i + C_p\mathbb{P}(X \leq \tau) = 1200 - 1000e^{-2\tau}$ and $ECL = \tau$. $g_2(\tau) = \frac{ECC}{ECL} = \frac{1200-1000e^{-2\tau}}{\tau}$.

(c) $g_2(\tau)$ has the following derivative: $\frac{d(g_2(\tau))}{d\tau} = \frac{\tau(2000e^{-2\tau})-(1200-1000e^{-2\tau})}{\tau^2} = \frac{(2000\tau+1000)e^{-2\tau}-1200}{\tau^2}$. $\frac{d(g_2(\tau))}{d\tau}$ is negative for all $\tau \leq 0.2$. Therefore, $\tau^* = 0.2$ years.

(d) Failure based is optimal due to constant failure rate of time to failure.

Solution 5.2. (a) $\mathbb{P}(Z_t = z) = p_z = e^{-\mu\tau}\frac{(\mu\tau)^z}{z!}, \quad z \in \mathbb{N}_0$.

(b) $\mathbb{P}(Z_t = z) = p_z = \binom{-\lambda\tau/\ln(1-q)+z-1}{z}(1-q)^{-\lambda\tau/\ln(1-q)}q^z, \quad z \in \mathbb{N}_0$

(c) $$p_{ij}^0 = \begin{cases} p_{j-i}, & \text{if } i \leq j < L; \\ p_{\geq L-i}, & \text{if } j = L; \\ 0, & \text{otherwise.} \end{cases}$$

(d) $p_{ij}^1 = p_j$

Solution 5.3. (a) Erlang distribution with shape parameter $k = 2$ and scale parameter $\lambda = 1/4$.

(b) No, because the remaining lifetime of the non-failed antenna will not improve by replacement due to lack of memory of the exponential distribution.

(c) Use DTM model with exponential time to defect with $C_i = 500$, $C_p = 3400$, $C_u = 2 \cdot 3400 + 11500 = 18300$ Euro. Let X denote the time to defect (lifetime of first antenna), Y the delay time (lifetime of second antenna), and $T = X + Y$. Now to compute the ECL, we need the following integral:

$$\begin{aligned} \int_0^\tau t f_T(t)dt &= \frac{1}{16}\int_0^\tau t^2 e^{-t/4}dt \\ &= \frac{1}{16}\left[-\frac{e^{-t/4}\left(\frac{1}{16}t^2+\frac{1}{2}t+2\right)}{\frac{1}{64}}\right]_{t=0}^{t=\tau} \\ &= \frac{1}{16}\left[-e^{-t/4}(4t^2+32t+128)\right]_{t=0}^{t=\tau} \\ &= 8-\left(\frac{1}{4}\tau^2+2\tau+8\right)e^{-\tau/4} \end{aligned} \tag{5.45}$$

so that

$$ECL = \int_0^\tau t f_T(t)dt + \tau(1-F_T(\tau)) = 8-\left(\frac{1}{4}\tau^2+2\tau+8\right)e^{-\tau/4} + \tau\left(e^{-\tau/4}+\frac{\tau}{4}e^{-\tau/4}\right). \tag{5.46}$$

The following integral will be needed to compute ECC:

$$\begin{aligned} \int_{x=0}^{\tau}(1-F_Y(\tau-x))f_X(x)dx &= \int_{x=0}^{\tau} e^{-(\tau-x)/4}\frac{1}{4}e^{-x/4}dx \\ &= \frac{1}{4}\int_{x=0}^{\tau} e^{-\tau/4}e^{x/4}e^{-x/4}dx \\ &= \frac{1}{4}\int_{x=0}^{\tau} e^{-\tau/4}dx \\ &= \frac{1}{4}\left[e^{-\tau/4}x\right]_{x=0}^{x=\tau} \\ &= \frac{\tau}{4}e^{-\tau/4}. \end{aligned} \tag{5.47}$$

(This last integral calculation can also be avoided completely by noticing that $\mathbb{P}(X < \tau \cap X + Y > \tau)$ is equivalent to the probability that in a Poisson process with arrival intensity $\frac{1}{4}$ exactly one arrival occurs in an interval of length τ so that the result can be found immediately from the Poisson distribution.) Now we have

$$ECC = 18300\left(1-e^{-\tau/4}-\frac{\tau}{4}e^{-\tau/4}\right)+3900e^{-\tau/4}\frac{\tau}{4}+500e^{-\tau/4} \tag{5.48}$$

Now the average cost rate is $g = ECC/ECL$.

(d) Numerical optimization yields $g^* = 1601.15$ Euro per year and $\tau^* = 1.50$ year.

(e) Option 3, because preventive replacement costs of 3400 Euro are only incurred if exactly one antenna fails, but the second does not. The conditioning suggested by Engineer 3 achieves this.

(f) The cost will be the exactly the same, because the age-based policy described is actually identical to the delay time model of this problem.

Solution 5.4. (a) Organizing transition probabilities in the same type of matrix as shown in Example 5.3 we find:

$$P^0 = \begin{pmatrix} 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.00000 & 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00039 \\ 0.00000 & 0.00000 & 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00336 \\ 0.00000 & 0.00000 & 0.00000 & 0.54881 & 0.32929 & 0.09879 & 0.02312 \\ 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.54881 & 0.32929 & 0.12190 \\ 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.54881 & 0.45119 \\ 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

(b) Similarly:

$$P^1 = \begin{pmatrix} 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \\ 0.54881 & 0.32929 & 0.09879 & 0.01976 & 0.00296 & 0.00036 & 0.00004 \end{pmatrix}.$$

(c) $M^* = 4$, $g^* = 679.92$ Euro per 12 weeks.

Solution 5.5. (a) Scale: $\hat{p} = 0.7401$. Shape: $\hat{r} = 3.6156$.

(b) $\lambda = 1.0884$, $q = 0.2599$.

(c)

$$P^0 = \begin{pmatrix} 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0004 \\ 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0011 \\ 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0032 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0090 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0248 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0645 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.1899 & 0.1569 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.3165 & 0.3468 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3367 & 0.6633 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

(d)

$$P^1 = \begin{pmatrix} 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \\ 0.3367 & 0.3165 & 0.1899 & 0.0924 & 0.0397 & 0.0157 & 0.0059 & 0.0021 & 0.0007 & 0.0002 & 0.0001 \end{pmatrix}$$

(e) -

(f)

$$P^1 = \begin{pmatrix} 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \\ 0.1684 & 0.3266 & 0.2532 & 0.1411 & 0.0661 & 0.0277 & 0.0108 & 0.0040 & 0.0014 & 0.0005 & 0.0003 \end{pmatrix}$$

(g) $g^* = 269.63$ Euro per week and $M^* = 6$.

Solution 5.6. (a) $R(x, a) = 2x - 350a$

(b) $V_{n+1}(x) = \max_{a \in \{0,1\}} \left(R(x, a) + \sum_{y=0}^{7000} p_{xy}^a V_n(y) \right)$

(c)

$$
\begin{aligned}
\max_{z_{x,a}} \quad & g = \sum_{x=0}^{7000} \sum_{a=0}^{1} R(x,a) z_{x,a} \\
\text{subject to} \quad & z_{x,0} + z_{x,1} = \sum_{y=0}^{7000} \sum_{a=0}^{1} z_{y,a} p_{yx}^{a}, \quad \forall x \in \{0, \dots, 7000\} \\
& \sum_{x=0}^{7000} \sum_{a=0}^{1} z_{x,a} = 1 \\
& z_{x,a} \geq 0, \quad \forall z \in \{0, \dots, 7000\}, \quad \forall a \in \{0, 1\}
\end{aligned} \tag{5.49}
$$

Chapter 6

Design of Maintenance Programs for Systems

> "Assumptions eat mathematical models for breakfast"
>
> Qiushi Zhu

Maintenance rules are rarely designed for components in isolation. Components belong to a technical system such as an aircraft, MRI-scanner, lithography machine, baggage handling system or building. Therefore, maintenance programs are usually designed for systems as a whole. The primary considerations that govern this design are avoiding down-time of the system, especially unscheduled down-time (USD) and cost of maintenance/replacement. A system usually needs to be down in order to replace or maintain any of its components. The consequence of down-time differ per capital asset, but in general they are quite severe and expensive. Possible consequences include the shut down of entire factories, delays of flights and other appointments with customers and all the ensuing costs and inconveniences. If we design a maintenance policy for each component in isolation, we can end up in a situation where we have to shut down the system very often to replace or inspect only one or two components. This is very disruptive for the operations that the system supports. Therefore, maintenance programs often have scheduled down-times, also called *scheduled downs*, during which it is possible do maintenance, inspections, and/or replacements for any component in the system. At these moments, we need to perform maintenance in such a way that the risk of unscheduled down-time due to failure before the next scheduled down-time is minimized. In principle, we then like to plan all inspections and preventive maintenance during these scheduled downs so as to minimize cost, while also minimizing unscheduled down-time between scheduled downs. However, a system consists of many components. For some components a form of condition based maintenance (CBM, see Chapter 5) is advisable, for others some form of usage based maintenance (UBM, see Chapter 3) and for some components, failures are completely random and failure based maintenance (FBM) is really the best one can do. CBM is usually only applied for expensive components whose condition can be measured with a sensor or a (simple) inspection. UBM is applied for components that degrade, but for which the condition is difficult or expensive to determine. Finally FBM is often applicable for electronics because electronics often have a constant failure rate. The goal of this chapter is to integrate some of the FBM, UBM and CBM models from Chapters 3 and 5 to design a maintenance program for an entire system.

6.1. Problem description

Consider a system consisting of multiple components that are subject to failure. Let I denote the set of all components in a system and let I_{CBM} denote the set of components for which we apply CBM, I_{FBM} denote the set of components for which we apply FBM, and I_{UBM} denote the set of components for which we apply UBM; $I_{CBM} \cap I_{UBM} = I_{FBM} \cap I_{UBM} = I_{CBM} \cap I_{FBM} = \emptyset$, $I_{FBM} \cup I_{CBM} \cup I_{UBM} = I$. The system will have a scheduled down every τ time units, so at times $0, \tau, 2\tau, \ldots$. In this chapter, τ is a decision variable. Each scheduled down incurs an expected cost of C_d so that the cost-rate due to scheduled downs is C_d/τ. C_d will include the cost for stopping a system and making technicians and equipment available to perform maintenance. For a production line in a factory, C_d will include costs for lost production time, costs for travel of maintenance technicians, rent of specialized equipment and possibly other costs. For a wind turbine park at sea, C_d will include the travel cost of the maintenance crew, rent of a vessel from which maintenance can be conducted and possibly other costs.

For each component $i \in I_{FBM}$, we apply failure based maintenance as explained in Section 3.2.1. We assume that the mean lifetime and corrective replacement cost $C_{u,i}$ are given. Section 3.2.1 explains how to evaluate the performance of a failure based policy for components $i \in I_{FBM}$. We let $c_{FBM,i}$ denote the average cost-rate incurred due to component $i \in I_{FBM}$.

For each component $i \in I_{UBM}$, the lifetime T_i, its distribution $F_{T_i}(t)$ and density $f_{T_i}(t)$ are given, as well as the cost of a minimal repair $C_{mr,i}$, corrective replacement $C_{u,i}$ and planned preventive replacement $C_{p,i}$. We apply the periodic usage based replacement policy with minimal repair as described in Section 3.2.5. The length of a period in this case is τ. For each component $i \in I_{UBM}$, we need to decide on the number of periods n_i before we apply preventive maintenance. The performance of such a policy can be evaluated using the results in Section 3.2.5. For fixed

τ, this section also explains how to optimize n_i. We let $c_{UBM,i}(n_i, \tau)$ denote the average cost-rate of a component $i \in I_{UBM}$ when scheduled downs are τ time units apart and n_i is used to decide after how many scheduled downs to replace preventively. Furthermore, we let $c^*_{UBM,i}(\tau) = c_{UBM,i}(n^*_i(\tau), \tau)$, where $n^*_i(\tau)$ minimizes $c_{UBM,i}(n_i, \tau)$ over $n_i \in \mathbb{N}$ for a given τ.

We partition I_{CBM} into two sets depending on whether components have delay time model degradation or Markovian degradation with a discrete state space. The first set is denoted I_{DTM} and the second set I_{MDP}.

Components $i \in I_{DTM}$ are assumed to have an exponential time to defect, X_i, and generally distributed delay time Y_i with corresponding known distributions and densities. Corrective maintenance costs $C_{u,i}$, preventive maintenance costs $C_{p,i}$, and minimal repairs cost $C_{mr,i}$ per repair are given. We control the system using the condition based maintenance model for delay time degradation in Section 5.1.3, where scheduled down times correspond to inspection epochs. The performance of such a policy for a given time between unscheduled downs can be computed using the results in Section 5.1.3. We let $c_{DTM,i}(\tau)$ denote the expected cost-rate incurred by component $i \in I_{DTM}$ when scheduled downs are τ time units apart. Note that we do not charge the inspection costs because they are included in the cost of a scheduled down C_d.

Components $i \in I_{MDP}$ are subject to Markovian degradation. We let $X_i(t)$ denote the degradation at time $t \in \mathbb{R}$ and (with a slight abuse of notation) $X_{i,t} = X_i(t\tau)$ with $t \in \mathbb{N}$. The state space of $X_{i,t}$ is $\mathcal{S}_i = \{0, \ldots, L_i\}$ where L_i denotes the failed state of component i. The transitions of $X_{i,t}$ depend on whether we apply replacement at a scheduled down. We let $p^a_{x,y,i}(\tau) = \mathbb{P}(X_{i,t+1} = y \mid X_{i,t} = x, A_{i,t} = a)$ where $A_{i,t}$ denotes the decision to replace component i at scheduled down t ($A_{i,t} = 1$ if we decide to replace, and 0 otherwise). Note also that we make the dependence of these transition probabilities on τ explicit because both τ and the replacement threshold will be decision variables in this context. The cost of a preventive replacement is $C_{p,i}$ the cost of a corrective replacement is now a function of τ for reasons explained in 5.2.3: $C_{u,i}(\tau) = C_{1,i}\tau + C_{2,i}$. We will assume that the degradation process $X_{i,t}$ is such that there exists an optimal control limit $M^*_i(\tau)$ that minimizes the cost-rate for a given time between scheduled downs τ. Finally, we let $c_{MDP,i}(M_i, \tau)$ denote the average cost-rate and $c^*_{MDP,i}(\tau) = c_{MDP,i}(M^*_i, \tau)$. These functions can be evaluated using the results in Section 5.2[1].

The optimization problem that we seek to solve is the following multi-variable non-linear non-

[1]Be careful here. The results in Section 5.2 describe how to compute the average cost per inspection epoch $g(M)$. Here we are interested in the average cost per time unit. This can be obtained by dividing the cost per inspection epoch over the length of the inspection interval/time between scheduled downs.

convex integer programming problem:

$$\min_{\tau,n_i,M_i} \sum_{i\in I_{FBM}} c_{FBM,i} + \sum_{i\in I_{UBM}} c_{UBM,i}(n_i,\tau) + \sum_{i\in I_{DTM}} c_{DTM,i}(\tau) + \sum_{i\in I_{MDP}} c_{MDP,i}(M_i,\tau) + \frac{C_d}{\tau}$$

$$\begin{aligned} \text{subject to} \quad & n_i \in \mathbb{N}, \qquad \forall i \in I_{UBM} \\ & M_i \in \mathbb{N}, \qquad \forall i \in I_{MDP} \\ & \tau \geq 0, \end{aligned} \tag{6.1}$$

which is equivalent to the following single variable non-linear non-convex programming problem:

$$\min_{\tau\geq 0} \quad C_{system} = \sum_{i\in I_{FBM}} c_{FBM,i} + \sum_{i\in I_{UBM}} c^*_{UBM,i}(\tau) + \sum_{i\in I_{DTM}} c_{DTM,i}(\tau) + \sum_{i\in I_{MDP}} c^*_{MDP,i}(\tau) + \frac{C_d}{\tau} \tag{6.2}$$

6.2. Solution approach

Optimization problems (6.1) and (6.2) are non-linear and non-convex, so we cannot hope for a very efficient algorithm. However, for a fixed τ, the objective of (6.2) can be approximately evaluated by assuming $c_{UBM,i}(n_i,\tau)$ are uni-modal in n_i for fixed τ and using a greedy search to find n_i^* for each $i \in I_{UBM}$. For components $i \in I_{MDP}$, we can determine $c^*_{MDP,i}(\tau)$ using the results in Section 5.2. A good solution for optimization problem (6.2) can therefore be found by evaluating $C_{system}(\tau)$ using the scheme above on a sufficiently fine grid. For example, one can evaluate $C_{system}(\tau)$ for $\tau = 0.01, 0.02, 0.03, \ldots$. Sometimes, there is a natural grid size because scheduled downs need to occur at known occasions such as weekends or the end of the month.

Example 6.1. Consider a Tunnel that consists of 4 components.

- Component 1 is an electronic system that monitors the number of vehicles in the tunnel. This electronic system has a constant failure rate of 0.05 failures per year.
- Component 2 is metal frame in the tunnel on which ventilation units are placed. The condition of this frame is difficult to asses, but the lifetime distribution is Weibul with shape $\beta = 5$ and scale $\eta = 2$.
- Component 3 is the ventilation unit in the tunnel. (See Figure 6.1.) The condition of this unit can be measured by an accelerometer. The maintenance engineer brings this accelerometer to the tunnel when it is shut down for inspection. The amplitude of vibration measured in m^{-4} degrades following a negative binomial process with $r = 2.2$ and $p = 0.15$. The ventilation unit is considered to have failed if the amplitude of vibration exceeds 60 m^{-5}.

- Component 4 is the surface concrete on which vehicles drive. The condition of the concrete is classified in 3 different states: (1) as good a new, (2) there are cracks but the width of these cracks does not exceed 1.0 cm, and (3) there are cracks and the width of cracks sometimes exceeds 1.0 cm. When the concrete is as good as new (state (1)), it takes an exponential amount of time with mean 5.5 years to move to state (2). Once in state (2), it takes an exponential amount of time with mean 1.0 year to move to failed state (3).

When one of these components needs to be replaced, the whole tunnel needs to be shut down as a safety precaution. This is only allowed to happen on weekends and it costs 9000 Euro to set-up these safety precautions as well as to compensate local businesses that suffer from the tunnel shut down.

When component 1 fails, it is usually fixed by inspecting various subcomponents and replacing the broken one. A specialized technician takes 4 hours to do this and charges 130 Euro per hour. The subcomponent that is replaced costs 50 Euro on average. These operations can be done without shutting down the tunnel.

Component 2 is maintained by a specialized contractor that charges 4500 Euro to restore the frame to as good as new condition. If the frame fails between inspection epochs, it has to be fixed temporarily (minimal repair) for 9100 Euro without closing down the tunnel. After such a minimal repair, the frame needs to be restored to as good as new condition for 4500 at the next planned down-time.

Component 3 is can be replaced by the original equipment manufacturer with an overhauled ventilation unit. This costs 8000 Euro if the ventilation unit has not failed yet. If the ventilation unit has already failed, a cost of 9500 Euro is charged to account for the extra cost to overhaul a failed ventilation unit. On top of this, the road safety inspection charges 700 Euro for every month since the last inspection if the ventilation unit is found to have failed during an inspection.

The surface concrete can be replaced entirely and this costs 7000 Euro. However, if all crack have a width of less then 1.0 cm, then the concrete does not need to be replaced and can be fixed by filling the cracks. Past experience indicates that this costs 3900 Euro. A minimal repair consists of locally filling a single crack and costs 3500 Euro. Inspecting the tunnel costs 50 Euro.

The situation above can be cast into the model of this chapter. In this case $C_d = 9000$ Euro, $I = \{1, 2, 3, 4\}$, $I_{FBM} = \{1\}$, $I_{UBM} = \{2\}$, $I_{MDP} = \{3\}$ and $I_{DTM} = \{4\}$. The cost per component can be determined as a function of the length of the scheduled down interval τ as follows:

- Component 1: Applying the model in Section 3.2.1 yields $c_{FBM,1} = 0.05 \cdot (4 \cdot 130 + 50) = 28.50$ Euro per year.

- Component 2: Applying the model in Section 3.2.5 we find can evaluate $c_{UBM,2}$. Note that

Figure 6.1: A ventilation unit

numerical integration is not needed for Weibul distributions.

- Component 3: Applying stochastic dynamic programming and using $C_{u,3}(\tau) = 9500 + 12 \cdot 700\tau$ the cost per scheduled down interval are determined. Dividing this by the length of the interval yields the cost per year.

- Component 4: Note that the model is a variation to Example 5.2 with different cost and lifetime parameters and follows the same procedure for evaluation.

The cost functions of components 1 to 4 as a function of τ (so with $n_2^*(\tau)$ and $M_3^*(\tau)$) are shown in Figure 6.2. Summing the cost of these four components and adding the term $C_d/\tau = 9000/\tau$ yields C_{system} as shown in Figure 6.3. The optimal scheduled down interval is $\tau^* = 3.2692$ year (which is 170 weeks) and the optimal annual cost $C^*_{system} = 10.68$ kEuro. ◇

6.2.1 Extensions

The model in this chapter can be extended in several ways. For surveys on so-called multi-item maintenance problems, we refer the reader to Nicolai and Dekker (2008), Cho and Parlar (1991), and Thomas (1986). Many extensions are based on approximations rather than exact models. These approximations capture some more realistic features of industrial problem such as *opportunistic maintenance*. When the system has an unscheduled down due to an unexpected failure, then this opportunity may be used to replace and or maintain several other components as well. See Zhu et al. (2014) for recent research on this topic.

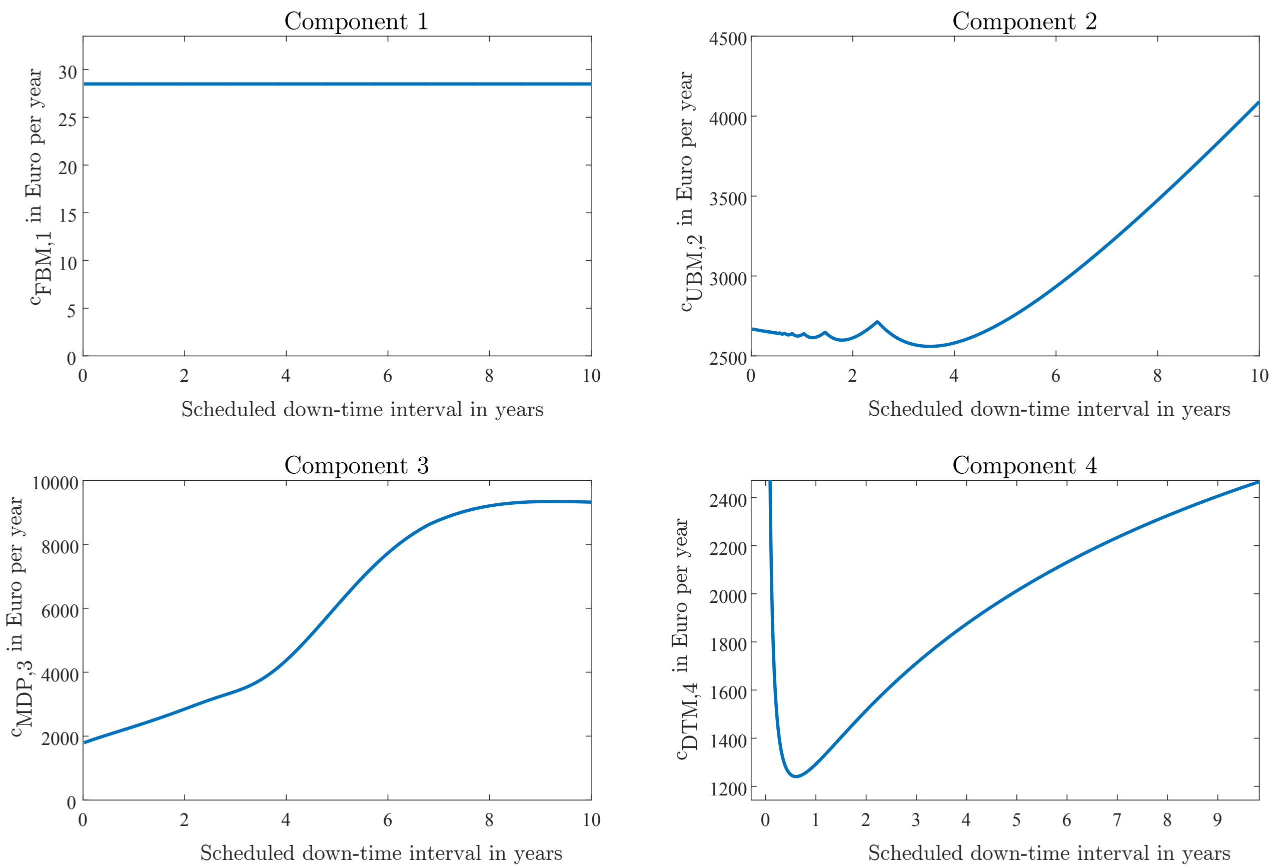


Figure 6.2: Annual cost per component vs scheduled down interval

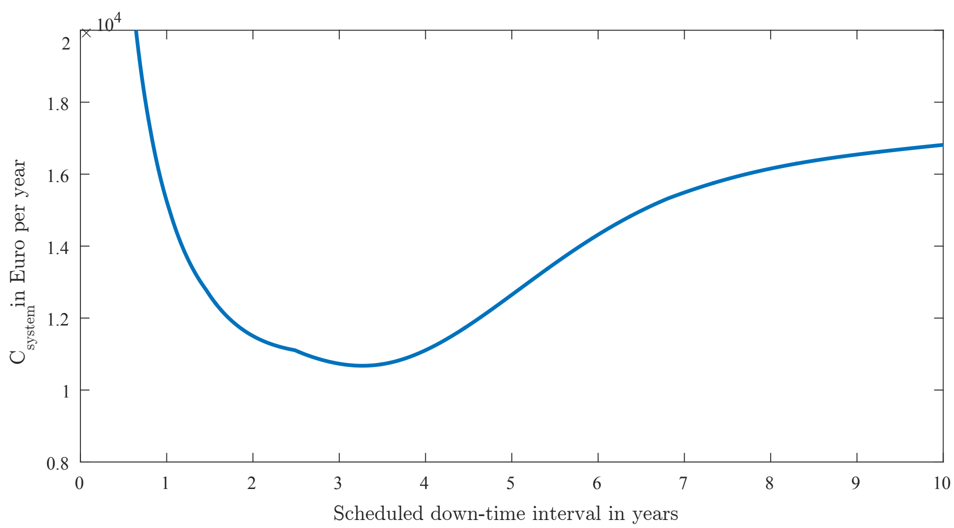


Figure 6.3: Total annual cost vs scheduled down interval

6.3. Exercises

Exercise 6.1. Replicate Example 6.1 by yourself.

Appendix A

Technical proofs

Lemma A.1. *Let N be a Poisson random variable with mean μ, D a logarithmic random variable with parameter p, and X a negative binomial random variable with shape r and scale $1-p$. Then their generating functions are given by*

$$\begin{aligned} G_N(z) &= \mathbb{E}[z^N] = \exp(\mu(z-1)), \quad z \in \mathbb{R} \\ G_D(z) &= \mathbb{E}[z^D] = \frac{\ln(1-pz)}{\ln(1-p)}, \quad |z| < \frac{1}{p} \\ G_X(z) &= \mathbb{E}[z^X] = \left(\frac{1-p}{1-pz}\right)^r, \quad |z| < \frac{1}{p} \end{aligned} \tag{A.1}$$

These results can be found in most good textbooks on probability so we omit the proof of this lemma.

Theorem A.1. *Consider a compound Poisson process with logarithmic compounding distribution. The logarithmic compounding distribution has parameter p and the corresponding random variables are denoted $(D_i)_{i\in\mathbb{N}}$. Furthermore, the Poisson arrival intensity is (without loss of generality) $-r\ln(1-p)$ and $N(t)$ denotes the number of Poisson arrivals in a period of length t. Then the compound Poisson process $X(t) = \sum_{i=1}^{N(t)} D_i$ has a negative binomial distribution with shape rt and scale $1-p$.*

Proof. We will show this result by showing that the generating function of $X(t)$ coincides with the generating function of a negative binomial random variable with said parameters.

$$\begin{aligned}
\mathbb{E}[z^{X(t)}] &= \mathbb{E}[z^{\sum_{i=1}^{N(t)} D_i}] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[z^{\sum_{i=1}^{N(t)} D_i} \mid N(t) = n\right] \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[z^{\sum_{i=1}^{n} D_i}\right] \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[z^{D}\right]^n \mathbb{P}(N(t) = n) \\
&= G_{N(t)}(G_D(z)) \\
&= \exp\left(-rt\ln(1-p)\left(\frac{\ln(1-pz)}{\ln(1-p)} - 1\right)\right) \\
&= \exp\left(-rt\left(\ln(1-pz) - \ln(1-p)\right)\right) \\
&= \exp\left(-rt\frac{\ln(1-pz)}{\ln(1-p)}\right) \\
&= \left(\frac{1-pz}{1-p}\right)^{-rt} \\
&= \left(\frac{1-p}{1-pz}\right)^{rt}
\end{aligned} \tag{A.2}$$

The fourth equality holds because the D_i are i.i.d. The fifth equality follows from the definition of the generating function. The sixth equality follows from using Lemma A.1 and the Poisson arrival intensity. □

Appendix B

Frequently occurring integrals

The following integrals will occur frequently in expressions that include the exponential distribution, Erlang distribution, and related distributions. It will be convenient to have these integrals ready for exercises and such. All of the integrals below can be found without too much effort by using integration by parts and/or substitution. However, they occur so frequently that it can be convenient to have them ready. In these expressions, C is an arbitrary constant.

$$\int e^{-ax} dx = -\frac{e^{-ax}}{a} + C \tag{B.1}$$

$$\int x e^{-ax} dx = -\frac{e^{-ax}(ax+1)}{a^2} + C \tag{B.2}$$

$$\int x^2 e^{-ax} dx = -\frac{e^{-ax}(a^2x^2+2ax+2)}{a^3} + C \tag{B.3}$$

$$\int x^3 e^{-ax} dx = -\frac{e^{-ax}(a^3x^3+3a^2x^2+6ax+6)}{a^4} + C \tag{B.4}$$

$$\int x^n e^{-ax} dx = -\frac{e^{-ax}\sum_{k=0}^{n} a^k x^k \frac{n!}{k!}}{a^{n+1}} + C, \quad n \in \mathbb{N}. \tag{B.5}$$

Bibliography

AberdeenGroup. Service parts management, unlocking value and profits in the service chain. *AberdeenGroup, Boston*, 2003.

Z. Benyamini and U. Yechiali. Optimality of control limit maintenance policies under non-stationary deterioration. *Probability in the Engineering and Informational Sciences*, 13:55–70, 1999.

D.I. Cho and M. Parlar. A survey of maintenance models for multi-unit systems. *European Journal of Operational Research*, 51(1):1 – 23, 1991. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/0377-2217(91)90141-H. URL `http://www.sciencedirect.com/science/article/pii/037722179190141H`.

J. Coetzee. *Maintenance*. Trafford publishing, 1997.

B.R. Ellingwood and Y. Mori. Probabilitic methods for condition assesment and life prediction of concrete structures in nuclear power plants. *Nuclear Engineering and Design*, 142:155–166, 1993.

A.H. Elwany and N.Z. Gebraeel. Sensor-driven prognostic models for equipment replacement and spare parts inventory. *IIE Transactions*, 40(7):629–639, 2008.

A.H. Elwany, N.Z. Gabraeel, and L.M. Maillart. Structured replacement policies for components with complex degradation processes and dedicated sensors. *Operations Research*, 59:684–695, 2011.

J.B. Jasper. Quick repsonse solutions: Fedex critical inventory logistics revitalized. *FedEx white paper*, 2006.

P. Kolesar. Minimum cost replacement under Markovian deterioration. *Management Science*, 12: 694–706, 1966.

C.J. Lu and W.Q. Meeker. Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 35:161–174, 1993.

D.C. Montgomery and G.C. Runger. *Applied statistics and probability for engineers*. Wiley, 4th edition, 2007.

R.P. Nicolai and R. Dekker. Optimal maintenance of multi-component systems: A review. In *Complex System Maintenance Handbook*, Springer Series in Reliability Engineering, pages 263–286. Springer London, 2008. ISBN 978-1-84800-010-0. doi: 10.1007/978-1-84800-011-7_11. URL http://dx.doi.org/10.1007/978-1-84800-011-7_11.

K.B. Öner, R. Franssen, G.P. Kiesmüller, and G.J. Van Houtum. Life cycle costs measurement of complex systems manufactured by an engineer-to-order company. In R.G. Qui, D.W. Russel, W.G. Sullivan, and M. Ahmad, editors, *The 17th International Conference on Flexible Automation and Intelligent Manufacturing*, pages 569–589, 2007.

M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley-Interscience, 1994.

S.M. Ross. *Stochastic Processes*. Wiley, 2 edition, 1996.

D. Stoneham. *Maintenance Management and Technology Handbook*. Elsevier advanced technology, 1998.

L.C. Thomas. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Engineering & System Safety*, 16:297–309, 1986.

L. Van Dongen. Maintenance engineering: Instandhouding van verbindingen. Oratie, UTwente, 2011.

J.M. van Noortwijk. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94:2–21, 2009.

M. Zhang, Z. Ye, and M. Xie. A condition-based maintenance strategy for heterogeneous populations. *Computers and Industrial Engineering*, 77:103–114, 2015.

Q. Zhu. *Maintenance Optimization of Multi-component Systems under Condition Monotoring*. PhD thesis, Eindhoven University of Technology, 2015.

Q. Zhu, H. Peng, and G.J. van Houtum. An opportunistic maintenance model for multi-component systems under a mixture of different maintenance policies. *BETA Working Paper, Eindhoven University of Technology*, 2014.