Control limit policies in a replacement model with additive phase-type distributed damage and linear restoration

David Perry

Department of Statistics, University of Haifa, Mt. Carmel, 31905 Haifa, Israel

Received 1 September 1998; received in revised form 1 January 2000

Abstract

We consider a replacement model for an additive damage process with linear restoration. The replacement policy is of the control limit type. Between replacements the damage level process is modeled as a compound Poisson process with positive jumps and negative drift, reflected at 0. We introduce meaningful cost functionals for this system and determine them explicitly for three special cases when the jumps have a phase-type distribution.

Keywords: Replacement model; Additive damage; Self restoration; Martingale stopping theorem; Phase-type distribution

1. Introduction

We consider a system working in a random environment that receives shocks at random points of time. Each shock causes a random amount of damage which accumulates over time. The damage process is controlled by means of a maintenance policy which causes the damage to decrease according to a piece-wise linear state-dependent restoration rate function $r(x) = 1$ for $x > 0$ and $r(0) = 0$. Thus, a breakdown can occur only upon the occurrence of a shock. Upon failure the system is replaced by a new identical one with a given cost. When the system is replaced before failure, a smaller cost is incurred, so that there is an incentive to attempt to replace the system before failure.

It was shown in [13] that for a model the optimal replacement policy is a so-called control limit level policy (CLL). That is, since record values occur only at times of shock arrivals there exists a certain critical level $\zeta$ such that the system is replaced once the damage process upcrosses it. Posner and Zuckerman [13] have shown that under natural assumptions the CLL policy is optimal for a large family of cost functions. However, they were only able to find the control limit by binary search. Perry and Posner [11], in a subsequent paper, restricted themselves to the case that the damage process is a regenerative process for which in each cycle...
the shocks are exponentially distributed and arrive in accordance with a compound Poisson process. Under these assumptions they extended the analysis to determine the optimal CLL for a certain cost function.

However, the assumptions made by Perry and Posner [11] are too restrictive. Firstly, their cost function includes only fixed, but not proportional maintenance costs. Secondly, they introduce explicit expressions for the cost functionals only for the case of exponential jumps.

This paper is an extension of Perry and Posner [11] in the sense that these two restrictive assumptions are removed. We add the proportional maintenance costs component to the cost function, and also, we make the first step towards generalization by assuming that the jump sizes have a phase-type distribution. Strictly speaking, we consider three special cases of phase-type distributions: exponential, Erlang (2, μ) and hyperexponential.

For the past three decades many researchers and practitioners have shown interest in the study of maintenance models for systems with stochastic failures using CLL policies (see Aven and Gaarder 1987, [20,23], the survey papers of Valdes-Flores and Feldman [21], Pierskalla and Voelker [12] and the references therein). One major reason for this is that maintenance models can be applied to a variety of stochastic models such as queues, inventories and insurance risk. For example, replacement can be interpreted as “disaster” or work removal by negative customers [4,7,8] in queueing models. It can be interpreted as an action of clearing in inventory models [14,18]. Under certain circumstances, it can be interpreted as a bankruptcy in insurance risk models [1, Chapter III]. Indeed, the damage level process, as defined in this study, is a regenerative process that can stay in zero for some time (this period of time is called an idle period). The risk process, in the insurance risk model is “killed” at the time of bankruptcy. The circumstances mentioned above represent the fact that the time to bankruptcy in the risk process can be interpreted as an arbitrary cycle whose idle period is deleted.

According to the CLL policy, the system is replaced once a certain level is upcrossed by the additive damage process. From a modeling point of view, our problem belongs to the family of first exit-time problems [2,3,5,6,15–17,24]. A key study in the area of additive damage is the paper by Taylor [19] in which he assumed that the damage process is a compound Poisson process. Aven and Gaarder (1987) extended his paper in a different direction. They also assumed that the cumulative damage is generated only by shock sizes; however, the system cannot restore the damage as time progresses.

In Section 2, we introduce the dynamics of the damage process and the cost functionals. Explicit expressions for the cost functionals are obtained in Section 3 for the exponential jump case, in Section 4 for the Erlang (2, μ) jump case and in Section 5 for the hyperexponential jump case. In Section 6, we introduce some numerical results and finally, in Section 7 we lay the ground for further research.

2. The model

Let \( W = \{W(t); t \geq 0\} \) be the additive damage level process. We assume that \( W \) is a Markov jump process in which the shocks arrive according to a Poisson process with rate \( \lambda \). Due to maintenance operation, we assume (without loss of generality) that the damage level process decreases at rate \( r(W(t)) = 1 \) whenever \( W(t) > 0 \) and \( r(0) = 0 \). A failure occurs once the damage level upcrosses a certain level \( \vartheta_0 \). Upon failure the system is replaced by a new identical one and the replacement cycles are repeated indefinitely. Each replacement costs \( C \) dollars and each failure adds an additional cost of \( K \) dollars. Thus, there is an incentive to attempt to replace the system before failure. We allow the controller to replace the system at any stopping time \( T \leq \delta \) where \( \delta \) is the failure time of the system. In addition, we assume proportional maintenance costs. It has been shown by Posner and Zuckerman [13] that the optimal replacement policy is that of the CLL. That is, there exists an optimal level \( 0 < \xi < \vartheta_0 \) such that whenever \( W \) upcrosses \( \xi \), the system is replaced and a new cycle starts, so that \( W \) is a regenerative process, and each cycle starts with damage level 0. We define the cycle length by \( T \) and note that failure may occur at the end of a cycle with a positive probability \( p \). Formally, let \( \{N(t); t \geq 0\} \) be a Poisson process with rate \( \lambda \) and define

\[
X(t) = V_1 + \cdots + V_{N(t)} - t,
\]  
(2.1)
where the $V_i$’s are i.i.d. random variables having Laplace transform (LT) $\phi(z)$. Define
\begin{equation}
I(t) = -\min_{0 \leq s \leq t} X(s)
\end{equation}
and
\begin{equation}
Z(t) = X(t) + I(t).
\end{equation}
Note that $I(t)$ is the amount of time until $t$ that the reflected process $Z(t)$ stays at 0. Define now the stopping time $T = \inf\{t > 0: Z(t) \geq \xi\}$. The damage level process $W = \{W(t): t \geq 0\}$ is a regenerative process whose cycle is a probabilistic replication of the stopped reflected process $\{Z(t): 0 \leq t \leq T\}$, so that the stopping time $T$ can be defined also with respect to $W$. Each cycle of $W$ starts from 0, $T$ is a point of discontinuity and $0 = W(T^+) \leq W(T^-) < W(T)$.

An upcrossing of level $\xi$ can also be an upcrossing of level $\theta_0(> \xi)$. As mentioned, this event occurs with probability $p$. That is, $p$ is the probability that the end of the cycle is the failure and $q = 1 - p$ is the probability for replacement before failure. Obviously, $p$ is a function of the decision variable $\xi$. Since $T$ is composed of a geometric number of cycles, it is clear that $E(T) = p$. In order to define the expected discounted costs over an infinite horizon, let $T_1 = T$ and $T_{n+1} = \inf\{t > T_n: W(t) \geq \xi\}$. That is, $T_n$ is the end of the $n$th cycle. Define the cost function
\begin{equation}
R_\xi = E\left[ \sum_{n=1}^{\infty} [C + K \cdot 1_{\{W(T_n) > \theta_0\}}] e^{-\beta T_n} + hE \int_0^{\infty} e^{-\beta t} W(t) \, dt, \right]
\end{equation}
where $\beta$ is the discount factor and $h dW(t)$ is the proportional maintenance cost per time unit if the damage level is $W(t)$ at the time interval $dt$.

Since $W$ is a regenerative process, $R_\xi$ can be trivially expressed in terms of three cycle cost functionals. This is done in the following lemma.

**Lemma 1.** Let
\begin{equation}
A_1(\xi) = E(e^{-\beta T}), \quad A_2(\xi) = E(e^{-\beta T} 1_{\{T=\delta\}})
\end{equation}
and
\begin{equation}
A_3(\xi) = E \left( \int_0^T e^{-\beta t} W(t) \, dt \right).
\end{equation}

Then
\begin{equation}
R_\xi = \frac{CA_1(\xi) + KA_2(\xi) + hA_3(\xi)}{1 - A_1(\xi)}.
\end{equation}

To derive the cost functionals of Lemma 1 we will use a martingale $\{M(t): t \geq 0\}$ that was introduced by Kella and Whitt [9] in another context (see also [10] for a Brownian inventory application). As will be seen, the martingale used by Perry and Posner [11] can be obtained as a special case of $M(t)$.

**Lemma 2.** For all $\beta \geq 0$, and $\alpha \geq 0$ the process
\begin{equation}
M(t) = [\alpha - \lambda(1 - \phi(\alpha)) - \beta] \int_0^t e^{-\alpha Z(s) - \beta s} \, ds + 1 - e^{-\alpha Z(t) - \beta t} - \alpha \int_0^t e^{-\beta s} \, dI(s)
\end{equation}
is a martingale.
Lemma 2 is restricted to the case that \( Z(t) \) is a reflected compound Poisson process, and we use the martingale for a simple first passage stopping time \( T \). Thus, \( EM(T) = EM(0) \) by the optional sampling theorem and since \( \{Z(t): t \leq T\} \) represents an arbitrary cycle of \( W \) we obtain
\[
[z - \lambda(1 - \phi(z)) - \beta]E \int_0^T e^{-\beta W(s)} \, ds = -1 + Ee^{-\beta W(T)} - \beta T + \alpha E \int_0^T e^{-\beta s} \, dI(s).
\]
(2.4)

To simplify the notation, let
\[
\phi(z, \beta) = z - \lambda(1 - \phi(z)) - \beta
\]
(2.5)
and
\[
\eta(\beta) = E \int_0^T e^{-\beta s} \, dI(s).
\]
(2.6)

Then, Eq. (2.4) can be written as
\[
\phi(z, \beta)E \int_0^T e^{-\beta W(s)} \, ds = -1 = Ee^{-\beta W(T)} - \beta T + \alpha \eta(\beta).
\]
(2.7)

3. Exponential jumps

The first application is to the case in which the shock sizes \( V_i \sim \exp(\mu) \). In this case \( \phi(z) = \mu/(\mu + z) \). Substituting this in (2.5) we obtain
\[
\phi(z, \beta) = z - \frac{\lambda z}{\mu + z} - \beta.
\]
(3.1)

Clearly, for \( \beta > 0 \), \( \phi(z, \beta) \) in (3.1) has two real roots \( z_1(\beta) > 0 > z_2(\beta) \) given by
\[
z_1(\beta), z_2(\beta) = \frac{-(\mu - \lambda - \beta) \pm \sqrt{(\mu - \lambda - \beta)^2 + 4\mu \beta}}{2}.
\]
(3.2)

Also, by the memoryless property, \( W(T) - \xi \sim \exp(\mu) \) and \( W(T) \) and \( T \) are independent. Thus,
\[
Ee^{-\beta W(T)} = e^{-\xi} \mu \frac{\mu}{\mu + z} E(e^{-\beta T}).
\]
(3.3)

By substituting the roots of (3.2) in the fundamental equation (2.7), the left-hand side of (2.7) vanishes and we obtain two equations with two unknowns:
\[
1 = e^{-\xi z_1(\beta)} \frac{\mu}{\mu + z_1(\beta)} E(e^{-\beta T}) + z_1(\beta) \eta(\beta), \quad 1 = e^{-\xi z_2(\beta)} \frac{\mu}{\mu + z_2(\beta)} E(e^{-\beta T}) + z_2(\beta) \eta(\beta).
\]
(3.4)

Solving for \( Ee^{-\beta T} \) and \( \eta(\beta) \) we obtain
\[
A_1(\xi) = E(e^{-\beta T}) = \frac{z_1(\beta) - z_2(\beta)}{z_2(\beta) - z_1(\beta)} \frac{\mu}{\mu + z_1(\beta) + z_2(\beta)(\mu/(\mu + z_1(\beta)))} \quad \text{and} \quad \eta(\beta) = \frac{e^{-\xi z_1(\beta)}(\mu/(\mu + z_1(\beta))) - e^{-\xi z_2(\beta)}(\mu/(\mu + z_2(\beta)))}{z_2(\beta) - z_1(\beta)}.
\]
(3.5)

Note that (3.5) is the cost functional \( A_1(\xi) \) of Lemma 1. Clearly, by the memoryless property,
\[
Ee^{-\beta T} 1_{[T=\delta]} = e^{-\mu \delta e^{-\xi}} Ee^{-\beta T},
\]
so that

\[ A_2(\xi) = e^{-\mu(\lambda - \dot{\xi})} A_1(\xi). \]  

(3.7)

To obtain \( A_3(\xi) \) divide both sides of (2.7) by \( \varphi(\alpha, \beta) \), take the derivative with respect to \( \alpha \) and set \( \alpha = 0 \).

We find that

\[ A_3(\xi) = \frac{(1 - (\lambda/\mu)) A_1(\xi) + \beta((1/\mu) + \xi - \eta(\beta))}{\beta^2}. \]  

(3.8)

The functionals \( A_1(\xi) \) and \( A_2(\xi) \) (but not \( A_3(\xi) \)) have already been obtained in [11].

4. Erlang jumps

Assume now that the shock sizes \( V_i \sim \text{Erl}(2, \mu) \) so that \( \varphi(\alpha, \beta) = (\mu/(\mu + \alpha))^2 \). Substituting in (2.5) we have

\[ \varphi(\alpha, \beta) = \alpha - \lambda - \frac{\mu}{\mu + \alpha} \frac{1}{(\mu + \alpha)^2} \beta. \]  

(4.1)

**Lemma 3.** For \( \beta \geq 0 \) the function \( \varphi(\alpha, \beta) \) has three real roots

\( \alpha_1(\beta) > 0 > \alpha_2(\beta) > -\mu > \alpha_3(\beta). \)

**Proof.** The three real roots of \( \varphi(\alpha, \beta) \) are the three match points of the function \((\mu/(\mu + \alpha))^2\) and the linear function \((1 + (\beta/\lambda) - (\alpha/\lambda))\), as seen in Fig. 1. \( \square \)

**Remark 1.** It is required that all terms of (2.7) are well-defined at the roots \( \alpha_i(\beta), \; i = 1, 2, 3 \). Apparently, \( \alpha_3(\beta) \) is obtained outside the domain of convergence of the LT \((\mu/(\mu + \alpha))^2\) since \( \alpha_3(\beta) < -\mu \). However, the function \( \varphi(\alpha, \beta) \) is analytic on \( \mathbb{C} \setminus \{-\mu\} \). The integral \( \int_0^T e^{-\varphi(\alpha, \beta) s} ds \) is easily seen to be an analytic function of \( \alpha \) for all \( \alpha \in \mathbb{C} \). Thus, after multiplication by \( \mu + \alpha \) Eq. (2.7) becomes an identity between analytic functions which holds for \( \alpha > -\mu \) and thus, by **analytic continuation**, for all \( \alpha \in \mathbb{C} \setminus \{-\mu\} \), in particular for \( \alpha = \alpha_i(\beta), \; i = 1, 2, 3 \). The same problem already occurs in the exponential case.

It is assumed that the shock sizes \( V_i \sim \text{Erl}(2, \mu) \), which is the convolution of two exponential \((\mu)\) phases. From the memoryless property either \( W(T - \xi) \sim \text{Erl}(2, \mu) \) if level \( \xi \) is upcrossed by the first phase of the
jump, or $W(T) - \xi \sim \exp(\mu)$ if $\xi$ is upcrossed by the second phase of the jump. Let $E_1$ ($E_2$) be the event that $\xi$ is upcrossed by the first (second) phase of the jump. Then

$$E e^{-zW(T) - \beta t} = e^{-z\xi} \left( \frac{\mu}{\mu + z} \right)^2 \theta_1(\beta) + e^{-z\xi} \frac{\mu}{\mu + z} \theta_2(\beta),$$

(4.2)

where

$$\theta_1(\beta) = E(e^{-\beta T} 1_{(E_1)})$$

(4.3)

and

$$\theta_2(\beta) = E(e^{-\beta T} 1_{(E_2)}).$$

(4.4)

Clearly, by the memoryless property the occurrence of $E_1$ implies that $W(t) - \xi \sim \text{Erl}(2, \mu)$ and $E_2$ implies that $W(t) - \xi \sim \exp(\mu)$.

Substituting $z = \alpha_i(\beta)$ on the left-hand side of (2.7) for $i = 1, 2, 3$, and then (4.2)–(4.4) on the right-hand side, we obtain three equations with three unknowns:

$$1 = a_i(\beta) \theta_1(\beta) + b_i(\beta) \theta_2(\beta) + \alpha_i(\beta) \eta(\beta), \quad i = 1, 2, 3,$$

(4.5)

where

$$a_i(\beta) = e^{-z_i(\beta)\xi} \left( \frac{\mu}{\mu + z_i(\beta)} \right)^2 \quad \text{and} \quad b_i(\beta) = e^{-z_i(\beta)\xi} \frac{\mu}{\mu + z_i(\beta)}.$$

Eq. (4.5) is a simple system of three linear equations whose solution is easily available via MATHEMATICA. However, the explicit formulas are lengthy and not very illuminating.

For the three cost functionals we have

$$A_1(\xi) = \theta_1(\beta) + \theta_2(\beta),$$

(4.6)

$$A_2(\xi) = e^{-(\theta_0 - \xi)}[\theta_1(\beta)(1 + \mu(\theta_0 - \xi)) + \theta_2(\beta)]$$

(4.7)

and

$$A_3(\xi) = [(\xi + (1/\mu)A_1(\xi) + (1/\mu)\theta_1(\beta) + \eta(\beta)]\beta - (1 - p)(1 - A_1(\xi)) \over \beta^2.$$}

(4.8)

**Remark 2.** According to (4.6), $A_1(\xi)$ is composed of the two cost functionals $\theta_1(\beta)$ and $\theta_2(\beta)$ for some fixed discount factor $\beta$. However, $\theta_1(\beta)$ and $\theta_2(\beta)$ are also improper LTs. Therefore, $\alpha_1(0)$ and $\alpha_2(0)$ are the probabilities to upcross level $\xi$ by either the first phase or the second phase of $W(T) - \xi$, respectively. Thus, by substituting $\beta = 0$ in (4.7) we obtain the probability $p$ that a replacement is also a failure.

## 5. Hyperexponential jumps

It is easy to show (although not obvious at first glance) that the simplest hyperexponential distribution (namely, the mixture between two exponential distributions) is also of phase-type (see [22, p. 269]). We have

$$\phi(\xi) = \pi \frac{\mu}{\mu + \xi} + (1 - \pi) \frac{\gamma}{\gamma + \xi}.$$

(5.1)

Namely, for some known $0 < \pi < 1$ the jump size is a mixture of two exponential random variables; one with parameter $\mu$ and one with parameter $\gamma$, and assume, without loss of generality, that $\mu > \gamma$. The function $\phi(\xi, \beta)$ for this case is different from that of Section 4 but Lemma 3 is still applicable in the sense that $\phi(\xi, \beta)$ has three real roots $\alpha_1(\beta) > 0 > \alpha_2(\beta) > -\gamma > \alpha_3(\beta) > -\mu$ for all $\beta \geq 0$. These three roots are the three match points between $\phi(\xi)$ of (5.1) and the linear function $1 + (\beta/\lambda) - (\xi/\lambda)$ as can be seen in Fig. 2.

To compute the functionals $A_i(\xi)$ we proceed step by step, exactly as in Section 4.
6. Numerical results

The whole analysis is carried out to find the optimal control level $\xi^*$. However, in practice the goal can be achieved only numerically. The expressions of the cost functionals, even in the case of exponential jumps, are too cumbersome for mathematical analysis. In this section we focus on the exponential case and provide some numerical results with respect to the objective function $R_\xi$ as given in Lemma 1. Starting with the base values of the parameters as

$$\lambda = 1, \quad \mu = 2, \quad \vartheta_0 = 8, \quad \beta = 0.9, \quad C = 10, \quad K = 15, \quad h = 3$$

we note first that $R_\xi$ is a convex function whose minimum is obtained at $\xi^* = 1.08$ and $R^*_{\xi} = 5.09$. In Table 1 below we have varied the parameters one at a time (shown in bold) to see the effect of this variation. For example, increasing the rate of arrival $\lambda$, as well as decreasing the value of $\mu$, results in an increase of $R_\xi$. However, $R_\xi$ is insensitive with respect to the parameters $\vartheta_0$ and $K$. There is a low sensibility with the parameters $\beta$ and $C$ and high sensibility with $h$. The changes of the parameters and the objective function $R_\xi$ are all in directions which agree with intuition.

7. A concluding remark

Asmussen and Perry [2] introduce a different approach to compute the LT of the first-exit time of a reflected Lévy process with phase-type jumps. We believe that the approach introduced in this paper may lead to more general results.
Indeed, the three applications presented in Sections 3–5 are the simplest cases of phase-type distributions. However, if we apply our approach to the case that the jump size \( \mathcal{E}(n; \mu) \), Eq. (2.7) should lead to \( n+1 \) equations, possibly generated by the roots of the function

\[
\phi(x, \beta) = x - \lambda \left( 1 - \left( \frac{\mu}{\mu + x} \right)^n \right) - \beta.
\]

The \( n+1 \) unknown functionals on the right-hand side of the martingale \( M(t) \) are \( \eta(\beta) \) and \( \theta_j(\beta) = E e^{\beta T} \mathbb{1}_{E_j} \), \( j = 1, \ldots, n \), where \( E_j \) designates the event that \( \zeta \) is upcrossed by the \( j \)th phase. That is, \( E_j \) implies \( W(T) - \zeta \sim \mathcal{E}(n - j + 1; \mu) \), so that

\[
E e^{-zW(T) - \beta T} = e^{-z\zeta} \sum_{j=1}^{n} \left( \frac{\mu}{\mu + x} \right)^j \theta_{n-j+1}(\zeta).
\]

However, in the case that the jump size \( \mathcal{E}(n; \mu) \) we have only two match points between the linear function \( 1 + (\beta/\lambda) - (x/\lambda) \) and the LT \( \phi(x) = (\mu/(\mu + x))^n \). These two match points are the two real roots at \( \phi(x, \beta) \). Thus, there must also be \( n - 1 \) extra complex roots. This problem is now under investigation.

References