EURANDOM PREPRINT SERIES

2020-005

April 23, 2020

Censored lifetime learning: Optimal Bayesian age-replacement policies

C. Drent, S. Kapodistria, O. Boxma
ISSN 1389-2355

Censored lifetime learning: Optimal Bayesian age-replacement policies

Collin Drent[a,*], Stella Kapodistria[a], Onno Boxma[a]

[a] *Eindhoven University of Technology, Department of Mathematics and Computer Science, P.O. Box 513, 5600 MB Eindhoven, the Netherlands*

Abstract

We consider a sequence of age-replacement problems with a general lifetime distribution parametrized by an a-priori unknown parameter. There is a trade-off: Preventive replacements are censored but cheap, whereas corrective replacements are uncensored but costly observations of the lifetime distribution. We first analyze the optimal policy for a finite sequence and establish some properties. We then propose a myopic Bayesian policy that almost surely learns the unknown parameter and converges to the optimal policy with full knowledge of the parameter.

Keywords: age-replacement, maintenance, Bayesian learning, censoring, asymptotic optimality

1. Introduction

In the classical age-replacement problem, introduced in Barlow and Hunter (1960), a decision maker determines the optimal age threshold to preventively replace a single-component system subject to random failures to avoid high costs and/or low reliability associated with corrective replacements. The key assumption in this canonical age-replacement problem, and in many of its variations (we refer the interested reader to Wang (2002), and De Jonge and Scarf (2019), as they provide comprehensive overviews of the area), is that the lifetime distribution is a-priori fully determined and known to the decision maker. However, in many real-life applications, especially when a component has not yet generated (sufficient) data to estimate the lifetime distribution, this assumption is unfounded and necessitates an age-replacement policy that integrates learning and decision making.

The concept of integrating learning in optimal decision making has recently gained momentum in the literature of condition-based maintenance (CBM). Elwany et al. (2011), Kim and Makis (2013), Chen et al. (2015), Van Oosterom et al. (2017), and Drent et al. (2020) all study optimal CBM policies in which the parameters of the degradation process of a component are only partially known to the decision maker. The parameters are then inferred based on the component's degradation observations (e.g., vibrations, temperature) using Bayesian learning, which leads to policies that outperform conventional approaches which do not integrate learning. Two crucial assumptions in this literature, when incorporating Bayesian learning, are that (i) the prior and the posterior degradation distribution belong to the same family, i.e. they are conjugate distributions; and (ii) the degradation observations are fully observable. These two assumptions combined imply that the posterior distributions remain tractable throughout the information accumulation process.

By contrast, in the study of age-replacement policies, the information accumulation process consists of both censored (i.e. preventive replacements) and uncensored (i.e. corrective replacements) observations of the underlying lifetime distribution. Consequently, for most distribution families, Bayesian updates lose their conjugate property and the problem becomes intractable. Hence, only few age-replacement papers have studied the integration of Bayesian learning in optimal decision making. The most relevant analyses (to the work of this paper) are Fox (1967) and Dayanik and Gürler (2002), who both consider a sequence of age-replacement problems, where the lifetime is a Weibull random variable. Fox (1967) assumes that only the scale parameter is unknown, and formulates a Bayesian dynamic program to analyze the optimal policy of an infinite sequence of age-replacement problems. The author shows that this Bayesian dynamic program converges to the corresponding dynamic program of the setting in which the scale parameter is known. The Bayesian dynamic program is, however, computationally intractable and therefore difficult to implement in practice. Dayanik and Gürler (2002) therefore propose a myopic Bayesian policy that, at least numerically, performs close to the setting in which there is full knowledge of the unknown parameters. The authors do not, however, establish whether the learning and/or decision making of this myopic policy converge to the setting with full knowledge.

Fortunately, Braden and Freimer (1991) introduced a class of distributions that preserve the conjugate property even under censoring: *newsboy distributions*. This class of distributions has received much attention from the inventory research community focused on inventory systems where demand in excess of the inventory level is lost and thus unobserved. Lariviere and Porteus (1999), Ding et al. (2002), Chen and Plambeck (2008), Lu et al. (2008), Bensoussan et al. (2009), Chen (2010), Bisi et al. (2011) and Mersereau (2015) all assume a newsboy distribution with unknown parameter, which permits an exact anal-

*Corresponding author

Email addresses: `c.drent@tue.nl` (Collin Drent), `s.kapodistria@tue.nl` (Stella Kapodistria), `o.j.boxma@tue.nl` (Onno Boxma)

ysis of the optimal policy under different variations of this lost sales inventory control problem with censored demand learning. Surprisingly, the parametric Bayesian framework for censored learning assumed in these papers has not found its way to the age-replacement community, although such inventory control problems share similarities with adaptive maintenance: e.g., the inventory level and the age-replacement threshold influence the information accumulation process of the unknown demand and lifetime distribution, respectively, in the same way.

The main contributions of this paper are as follows. This paper is the first to investigate the effect of lifetime censoring on the optimal policy in a sequence of age-replacement problems where the lifetime distribution can be expressed as a newsboy distribution. We first analyze the optimal policy for a finite sequence of components, where there is an inherent exploration-exploitation trade-off. As this optimal policy is analytically intractable, we then propose a computationally appealing Bayesian policy. We show that this policy is asymptotically optimal; that is, (i) it almost surely learns the unknown parameter, and (ii) it converges to the optimal decision one would have taken with full knowledge of the unknown parameter.

The remainder of this paper is organized as follows: We introduce our notation and the problem formulation in §2. In §3, we provide some background on censored learning in the parametric Bayesian framework for newsboy distributions, and establish a new stochastic order result for this class of distributions. We analyze the optimal policy when the sequence of components is finite in §4. In §5, we propose a myopic Bayesian policy and establish its asymptotic properties. Finally, §6 contains some concluding remarks.

2. Problem formulation

We consider a sequence of N components, where the components are indexed by $n = 1, 2, \ldots, N$ defined over a probability space $(\Omega, \mathcal{F}, P)$. Each component is controlled by the classical age-replacement policy. That is, component n is either replaced preventively when it has been in operation for a fixed amount of time τ_n (at cost $c_p > 0$), or it is replaced correctively if it fails before this time (at cost $c_c > c_p$).

We assume that the lifetimes of the components are independent and identically distributed, belonging to a family of distributions parametrized by an unknown parameter $\theta > 0$ with true value θ_0. Given parameter θ, the lifetime distribution has a probability density function denoted by $f_X(x \,|\, \theta)$ and a cumulative distribution function denoted by $F_X(x \,|\, \theta)$. To preserve the conjugate property under censored learning, we assume that the underlying lifetime distribution is from the class of newsboy distributions, so that $F_X(x \,|\, \theta)$ can be expressed in the form

$$F_X(x \,|\, \theta) = 1 - e^{-\theta \ell(x)}, \tag{1}$$

where $\ell(x) : [0, \infty) \to [0, \infty)$ is a differentiable, nondecreasing and unbounded function with $\ell(0) = 0$ that is known to the decision maker (Braden and Freimer 1991). Note e.g., that the Weibull distribution with known shape parameter $\beta > 0$ and unknown scale parameter θ can be expressed as a newsboy distribution by setting $\ell(x) := x^{\beta}$. The Weibull distribution has been extensively used in modeling lifetimes due to its ability to model various aging classes of lifetime distributions (cf. Ahmad and Kamaruddin 2012). For a further detailed discussion on this broad class of distributions, we refer the interested reader to Braden and Freimer (1991).

The decision maker can only observe censored observations, rather than lifetime realizations. For component n, the censored observation is given by $X_n \wedge \tau_n := \min\{X_n, \tau_n\}$, where X_n is the realized lifetime and τ_n the imposed age-replacement threshold. Let $\mathcal{F}_n$ be the filtration generated by this censored lifetime process. Hence, for $n \geq 1$, we have the σ-algebra

$$\mathcal{F}_n := \sigma(X_1 \wedge \tau_1, \tau_1, X_2 \wedge \tau_2, \tau_2, \ldots, X_n \wedge \tau_n, \tau_n),$$

and let $\mathcal{F}_0$ be the trivial σ-algebra. It is thus evident that the accumulated information about the lifetime up to the $n-$th component is impacted by all past replacements decisions.

The decision maker wishes to minimize the total expected discounted cost due to both corrective and preventive replacements, where costs are discounted with rate $\alpha \in (0, 1)$, over an N component horizon. This optimality criterion is often employed in finite-horizon problems in the age-replacement literature (cf. De Jonge and Scarf 2019). We are interested in finding the optimal non-anticipatory policy (i.e. the decision τ_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$) that attains this minimum.

Observe that when deciding on the age-replacement threshold for the $n-$th component, there is an inherent trade-off between the direct expected cost of the $n-$th component and the impact the decision has on future costs through the information accumulation process. Specifically, exploration (i.e. a higher value of the age-replacement threshold) increases the probability of a corrective replacement, but at the same time leads to accumulating more, valuable, information. This phenomenon is often referred to as the exploration-exploitation trade-off.

3. Censored lifetime learning

In this section, we describe how the lifetime distribution of components can be inferred with increasing accuracy as information is accumulated. The approach is based on Braden and Freimer (1991), who show that the gamma distribution is the conjugate prior for all newsboy distributions.

Following a Bayesian approach, we treat the unknown parameter as a random variable, denoted with Θ, and assume that the decision maker has a prior density for the unknown parameter θ, denoted by $p_\Theta(\theta)$. This density captures the information about the unknown parameter of the lifetime distribution.

Let m and k denote the shape and scale parameter, respectively, of the gamma distribution. The prior density is then given by

$$p_\Theta(\theta \,|\, m, k) := \frac{k^m \theta^{m-1} e^{-k\theta}}{\Gamma(m)}, \quad \text{for all } \theta > 0, \tag{2}$$

where $\Gamma(\cdot)$ denotes the gamma function. Using $f_X(x \,|\, \theta) = \frac{d}{dx} F_X(x \,|\, \theta) = \theta \ell'(x) e^{-\theta \ell(x)}$ (cf. Equation (1)) and unconditioning on θ using Equation (2), we obtain the posterior predictive lifetime density and distribution function, respectively;

$$f_X(x|m,k) = \frac{mk^m \ell'(x)}{[k+\ell(x)]^{m+1}} \text{ and } F_X(x|m,k) = 1 - \left[\frac{k}{k+\ell(x)}\right]^m.$$

We use the shorthand notation m_n and k_n to denote the updated shape and scale parameter conditional on $\mathcal{F}_n$. Here, we omit the dependence on $\mathcal{F}_n$ as there is a mapping between $\mathcal{F}_n$ and (m_n, k_n), which we explain henceforth. Let (m_0, k_0) denote the parameters before the installment of the first component. Then for $n \geq 1$, conditional on $\mathcal{F}_n$, the prior hyper parameters are computed as:

$$\begin{aligned} m_n &= m_0 + \sum_{i=1}^{n} \mathbb{1}_{\{X_i < \tau_i\}} = m_{n-1} + \mathbb{1}_{\{X_n < \tau_n\}}, \text{ and,} \\ k_n &= k_0 + \sum_{i=1}^{n} \ell(X_i \wedge \tau_i) = k_{n-1} + \ell(X_n \wedge \tau_n), \end{aligned} \tag{3}$$

where $\mathbb{1}_{\{a\}}$ denotes the indicator function taking value 1 if event a occurs and value 0 otherwise. Here, we assume that $m_0 \cdot k_0 > 1$, and thus $m_n \cdot k_n > 1$ for all $n \geq 1$ through the update rules, so that the posterior predictive lifetime distribution of each component has a finite expectation. Observe that the scale parameter k_n is an aggregate of all observations. The shape parameter m_n counts the number of uncensored observations (corresponding to corrective replacements), and, as the coefficient of variation for the gamma prior is equal to $\sqrt{1/m_n}$, it is also a measure for the precision of the accrued information on the unknown parameter θ.

Equation (3) induces a simple, Markovian scheme for sequentially inferring the lifetime distribution conditional on $\mathcal{F}_n$.

In what follows, for notational simplicity and in order to enhance the readability of the paper, we write, depending on our objective, either $\mathcal{F}_n$ or (m_n, k_n).

In order to derive structural properties of the conditional posterior predictive lifetime random variable given (m_n, k_n), denoted with $X(m_n, k_n) := \{X \mid m_n, k_n\}$, and in order to make comparisons between different conditional posterior predictive lifetime random variables, we use the hazard rate ordering:

Definition 1 (Shaked and Shanthikumar 2007)**.** Let Y and Z be two nonnegative random variables with absolutely continuous distribution functions and with hazard rate functions $r(x)$ and $q(x)$, respectively, such that $r(x) \geq q(x)$ for all $x > 0$. Then Y is said to be smaller than Z in the hazard rate order.

We now present an important proposition that indicates how the accrued information, encoded in (m, k), affects the stochastic ordering of the conditional posterior predictive lifetime distribution.

Proposition 1. *The conditional posterior predictive lifetime random variable $X(m, k)$ is:*
(i) stochastically increasing in the hazard rate order in the scale parameter k, and
(ii) stochastically decreasing in the hazard rate order in the shape parameter m.

Proof. Let $h(x \mid m, k)$ denote the hazard rate function of the conditional posterior predictive lifetime when the shape and scale parameters are m and k, respectively. We then have

$$h(x \mid m,k) := \frac{f_X(x \mid m,k)}{1 - F_X(x \mid m,k)} = \frac{\frac{mk^m \ell'(x)}{[k+\ell(x)]^{m+1}}}{\left[\frac{k}{k+\ell(x)}\right]^m} = \frac{m\ell'(x)}{k+\ell(x)}.$$

Observe that since $\ell(x)$ is positive and increasing (by assumption), we have the following. If $M > m > 0$ then $h(x \mid M, k) > h(x \mid m, k)$ for all $x > 0$, which establishes the hazard rate order in m. Finally, if $K > k > 0$ then $h(x \mid m, K) < h(x \mid m, k)$ for all $x > 0$, which establishes the hazard rate order in k. □

Assertion (i) establishes the monotonic increase in the expected conditional posterior predictive lifetime when the aggregate of observations increases for a fixed number of uncensored observations. This implies that if the aggregate of all observations is high, then past components have had a relatively long lifetime on average. Hence, the decision maker predicts that the next component will have a longer lifetime than in the case where the accrued information has a lower aggregate of observations.

Assertion (ii) establishes the stochastic-ordering property of lifetime distributions that are updated using censored observations and uncensored observations, respectively. It states that a censored lifetime observation results in a lifetime distribution that is stochastically greater than that from an uncensored observation. An intuitive explanation is as follows. With a censored observation, the true lifetime is at least as large as the censored observation, as opposed to the uncensored observation, where the true lifetime is equal to the uncensored observation.

As the usual stochastic order is implied by the hazard rate order (see e.g., Shaked and Shanthikumar 2007, Theorem 1.B.1.), Proposition 6.2 and 6.3 of Braden and Freimer (1991), which state the usual stochastic order of the conditional posterior predictive lifetime, directly follow from Proposition 1.

4. Optimal policy for a finite sequence

In this section, we investigate the structure of the optimal policy when N is finite. In a finite sequence of components, the exploration-exploitation trade-off urges the decision maker to explicitly recognize the impact that current decisions have on both the direct expected costs and the future expected costs through the information accumulation process. This interdependence is made explicit by formulating the optimization problem as a dynamic program.

To this end, let $V_n(m, k)$ denote the minimum total expected discounted cost over components $n, n+1, \ldots, N$, starting with component n, when the updated hyper parameters are (m, k), respectively. We assume the terminal cost to be zero, hence $V_{N+1}(m, k) := 0$ for all (m, k). The optimality equations, for $n = 1, 2, \ldots, N$, are:

$$V_n(m,k) = \min_{\tau_n \geq 0} \Big\{ C_n(\tau_n \mid m,k) + \int_0^\infty G_n(\tau_n, x \mid m,k) f_X(x \mid m,k) \mathrm{d}x \Big\}, \tag{4}$$

where

$$C_n(\tau_n\,|\,m,k) = \tag{5}$$
$$c_c\int_0^{\tau_n} e^{-\alpha x} f_X(x\,|\,m,k)\mathrm{d}x + c_p e^{-\alpha\tau_n}(1 - F_X(\tau_n\,|\,m,k)),$$

denotes the direct expected discounted cost function of component n when the age-replacement threshold is τ_n (the first part is due to corrective replacement and the second part is due to preventive replacement) and

$$G_n(\tau_n, x\,|\,m,k) := \begin{cases} e^{-\alpha x}V_{n+1}(m+1, k+\ell(x)), & \text{if } x < \tau_n, \\ e^{-\alpha\tau_n}V_{n+1}(m, k+\ell(\tau_n)), & \text{if } x \geq \tau_n, \end{cases} \tag{6}$$

denotes the discounted cost function over the remaining components $n+1,\ldots,N$ when the age-replacement threshold of the n−th component is τ_n and the lifetime realization equals x.

The existence of an optimal policy in this setting is guaranteed, see e.g., Proposition 3.4 of Bertsekas and Shreve (1978). Observe that over an N component horizon, the minimum total expected discounted cost is given by $V_1(m_0, k_0)$ which can be found by solving Equation (4) via backward induction. Unfortunately, analytic solutions do not appear to be readily available and solving Equation (4) numerically, even for small instances, is a difficult task. Instead, we establish some structural results of the optimal policy in the remainder of this section and then focus on an asymptotically optimal policy that is simple to compute making it easy to implement in practice.

We first state two properties regarding the direct expected discounted cost function that are instrumental in characterizing the behavior of $V_n(m,k)$ with respect to its parameters.

Lemma 1. *$C_n(\tau_n\,|\,m,k)$ is*
(i) nonincreasing in k, and
(ii) nondecreasing in m,
for all $\tau_n \geq 0$ and $n \in \{1,2,\ldots,N\}$.

Proof. Note that Equation (5) can be rewritten as

$$C_n(\tau_n\,|\,m,k) =$$
$$-c_c\int_0^{\tau_n} e^{-\alpha x}\mathrm{d}((1 - F_X(x\,|\,m,k)) + c_p e^{-\alpha\tau_n}(1 - F_X(\tau_n\,|\,m,k))$$
$$= (c_p - c_c)e^{-\alpha\tau_n}(1 - F_X(\tau_n\,|\,m,k)) + c_c$$
$$-c_c\alpha\int_0^{\tau_n} e^{-\alpha x}(1 - F_X(\tau_n\,|\,m,k))\mathrm{d}x. \tag{7}$$

Since $c_p < c_c$, $\alpha \in (0,1)$, and because of Assertion (i) (Assertion (ii)) of Proposition 1, both terms involving $F_X(\tau\,|\,m,k)$ in the last two lines of (7) are nonincreasing (nondecreasing) in k (m), which establishes the result. □

We now proceed with two properties regarding the minimum total expected discounted cost over components $n, n+1,\ldots,N$, starting with the n−th component.

Theorem 1. *$V_n(m,k)$ is*
(i) nonincreasing in k, and
(ii) nondecreasing in m,
for all $n \in \{1,2,\ldots,N+1\}$.

Proof. We first prove Assertion (i) by backward induction. The base case, i.e. $V_{N+1}(m,k)$, holds trivially as terminal costs $V_{N+1}(m,k) = 0$ for all (m,k). Let $K > k$ and assume inductively that $V_{n+1}(m,K) \leq V_{n+1}(m,k)$, and let $\tau^n_{m,j}$ denote the optimal age-replacement threshold for the n−th component, when the shape and scale parameters are m and j, respectively. We have

$$\begin{aligned}
&V_n(m,K) - V_n(m,k) \\
&= C_n(\tau^n_{m,K}\,|\,m,K) + \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,K)f_X(x\,|\,m,K)\mathrm{d}x \\
&\quad - C_n(\tau^n_{m,k}\,|\,m,k) - \int_0^\infty G_n(\tau^n_{m,k}, x\,|\,m,k)f_X(x\,|\,m,k)\mathrm{d}x \\
&\leq C_n(\tau^n_{m,K}\,|\,m,K) + \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,K)f_X(x\,|\,m,K)\mathrm{d}x \\
&\quad - C_n(\tau^n_{m,K}\,|\,m,k) - \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,k)f_X(x\,|\,m,k)\mathrm{d}x \\
&\leq \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,K)f_X(x\,|\,m,K)\mathrm{d}x \\
&\quad - \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,k)f_X(x\,|\,m,k)\mathrm{d}x \\
&\leq \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,k)f_X(x\,|\,m,K)\mathrm{d}x \\
&\quad - \int_0^\infty G_n(\tau^n_{m,K}, x\,|\,m,k)f_X(x\,|\,m,k)\mathrm{d}x \\
&= \mathbb{E}[G_n(\tau^n_{m,K}, X\,|\,m,k)\,|\,m,K] - \mathbb{E}[G_n(\tau^n_{m,K}, X\,|\,m,k)\,|\,m,k] \leq 0.
\end{aligned}$$

The first inequality holds because $\tau^n_{m,K}$ is a feasible policy for m and k but not necessarily optimal. The second inequality holds by Assertion (i) of Lemma 1. The third inequality follows from the induction hypothesis. That is, $V_{n+1}(m,K) \leq V_{n+1}(m,k)$ implies that

$$\begin{aligned}
&G_n(\tau^n_{m,K}, x\,|\,m,K) - G_n(\tau^n_{m,K}, x\,|\,m,k) \\
&= \begin{cases} e^{-\alpha x}(V_{n+1}(m+1, K+\ell(x)) - V_{n+1}(m+1, k+\ell(x))) \leq 0, \\ e^{-\alpha\tau^n_{m,K}}(V_{n+1}(m, K+\ell(\tau^n_{m,K})) - V_{n+1}(m, k+\ell(\tau^n_{m,K}))) \leq 0, \end{cases}
\end{aligned}$$

where the first branch corresponds to $x < \tau^n_{m,K}$ and the second branch to $x \geq \tau^n_{m,K}$. Hence, $G_n(\tau^n_{m,K}, x\,|\,m,K) \leq G_n(\tau^n_{m,K}, x\,|\,m,k)$ for all $x \geq 0$, which implies the third inequality. Then, following a similar reasoning, $G_n(\tau^n_{m,K}, x\,|\,m,k)$ is nonincreasing in x by the induction hypothesis and since $e^{-\alpha x}$ is decreasing in x for $\alpha > 0$. We then have the expectation of a decreasing function $G_n(\tau^n_{m,K}, x\,|\,m,k)$, so that the last inequality is implied by the stochastic order of Assertion (i) of Proposition 1 between $f_X(x\,|\,m,K)$ and $f_X(x\,|\,m,k)$ (see e.g., Ross 1996, Proposition 9.1.2).

The proof of Assertion (ii) follows verbatim the proof of Assertion (i), starting with $M > m$, looking at the difference $V_n(m,k) - V_n(M,k)$, and assuming inductively that $V_{n+1}(m,k) \leq V_{n+1}(M,k)$ with $M > m$. □

Theorem 1 establishes the monotonicity of the minimum total expected discounted cost in both the aggregate of all observations k and the number of exact observations m. The intuition behind both parts (and their individual counterparts in Lemma 1) is as follows: If the aggregate of all observations increases and everything else is held fixed, it means that on average, each component has had a longer lifetime and is thus discounted at a higher rate. This leads to a lower total expected discounted cost. A similar reasoning holds if the number of exact observations increases and the aggregate of all observations is held fixed. Both preventive and corrective replacement costs are then, on average, discounted at an equal rate, but there are more exact observations so that c_c is incurred more often. This leads to a higher total expected discounted cost.

5. An asymptotically optimal policy

In the previous section we have established some structural results pertaining to the optimal policy in the case of a finite sequence of components (i.e. $N < \infty$). However, computing this optimal policy via the proposed dynamic program is analytically intractable and even numerically a difficult task. Therefore, in this section, we investigate asymptotic properties (as $N \to \infty$) of the information accumulation process and of a computationally tractable myopic policy.

We proceed in two steps. First we show that under any reasonable policy, the learning converges in the Bayesian sense. We then propose a myopic policy and prove its asymptotical optimality.

5.1. Convergence of learning

Recall that $\{\Theta \,|\, \mathcal{F}_n\}_{n=1,2,\dots,N}$ denotes our posterior belief regarding the unknown parameter θ and that the hyper parameters are updated according to Equation (3). In the next result, we show that this posterior expectation, denoted with $\mathbb{E}[\Theta \,|\, \mathcal{F}_N]$ converges (a.s.) to the true value θ_0, as $N \to \infty$, and that the variance, denoted with $\mathbb{V}ar[\Theta \,|\, \mathcal{F}_N]$, converges (a.s.) to 0, as $N \to \infty$. This convergence is guaranteed under any policy in which the threshold can be lower bounded by some constant $\epsilon > 0$. We believe that this assumption is justified as it is not natural to replace a component immediately after installment.

Theorem 2. *Under any policy for which $\tau_n \geq \epsilon > 0$ for all $n \in \{1, 2, \dots, N\}$, we have*

$$\mathbb{E}[\Theta \,|\, \mathcal{F}_N] \xrightarrow{a.s.} \theta_0 \text{ and } \mathbb{V}ar[\Theta \,|\, \mathcal{F}_N] \xrightarrow{a.s.} 0 \text{ as } N \to \infty.$$

Proof. By the updating scheme of the posterior density, the expectation of Θ given $\mathcal{F}_N$ and fixed policy $\tau_n \geq \epsilon > 0$ for all $n \in \{1, 2, \dots, N\}$ can be written as (recall that the random variable $\{\Theta \,|\, \mathcal{F}_N\}$ is gamma distributed, cf. Equation (2)):

$$\mathbb{E}[\Theta \,|\, \mathcal{F}_N] = \frac{m_N}{k_N} = \frac{m_0 + \sum_{n=1}^N \mathbb{1}_{\{X_n < \epsilon\}}}{k_0 + \sum_{n=1}^N \ell(X_n \wedge \epsilon)} = \frac{\frac{m_0}{N} + \frac{\sum_{n=1}^N \mathbb{1}_{\{X_n < \epsilon\}}}{N}}{\frac{k_0}{N} + \frac{\sum_{n=1}^N \ell(X_n \wedge \epsilon)}{N}}.$$

Then, when $N \to \infty$, we have by the strong law of large numbers, almost surely,

$$\lim_{N\to\infty} \mathbb{E}[\Theta \,|\, \mathcal{F}_N] = \lim_{N\to\infty} \frac{\frac{m_0}{N} + \frac{\sum_{n=1}^N \mathbb{1}_{\{X_n < \epsilon\}}}{N}}{\frac{k_0}{N} + \frac{\sum_{n=1}^N \ell(X_n \wedge \epsilon)}{N}} = \frac{F_X(\epsilon \,|\, \theta_0)}{\mathbb{E}[\ell(X \wedge \epsilon)]}.$$

Using straightforward calculus yields

$$\begin{aligned}
\frac{F_X(\epsilon \,|\, \theta_0)}{\mathbb{E}[\ell(X \wedge \epsilon)]} &= \frac{F_X(\epsilon \,|\, \theta_0)}{\int_0^\infty \mathbb{P}(\ell(X \wedge \epsilon) > y)\mathrm{d}y} \\
&= \frac{F_X(\epsilon \,|\, \theta_0)}{\int_0^{\ell(\epsilon)} \mathbb{P}(X > \ell^{-1}(y))\mathrm{d}y} \\
&= \frac{F_X(\epsilon \,|\, \theta_0)}{\int_0^{\ell(\epsilon)} e^{-\theta_0 y}\mathrm{d}y} \\
&= \frac{F_X(\epsilon \,|\, \theta_0)}{\frac{1}{\theta_0} F_X(\epsilon \,|\, \theta_0)} \\
&= \theta_0.
\end{aligned}$$

The second equality follows from the nondecreasing property of the function $\ell(\cdot)$. The third and fourth equality follow using the cumulative distribution function of a newsboy distribution, see Equation (1).

For the second part, note that

$$\begin{aligned}
\mathbb{V}\text{ar}[\Theta \,|\, \mathcal{F}_N] &= \frac{m_0 + \sum_{n=1}^N \mathbb{1}_{\{X_n < \epsilon\}}}{\left(k_0 + \sum_{n=1}^N \ell(X_n \wedge \epsilon)\right)^2} \\
&= \mathbb{E}[\Theta \,|\, \mathcal{F}_N] \cdot \frac{1}{k_0 + \sum_{n=1}^N \ell(X_n \wedge \epsilon)}.
\end{aligned}$$

Using

$$\mathbb{E}[\Theta \,|\, \mathcal{F}_N] \xrightarrow{a.s.} \theta_0 \text{ and } \frac{1}{k_0 + \sum_{n=1}^N \ell(X_n \wedge \epsilon)} \xrightarrow{a.s.} 0,$$

when $N \to \infty$ leads to the desired result. □

Observe that the crucial part of the proof, i.e. the equality $\frac{F_X(\epsilon \,|\, \theta_0)}{\mathbb{E}[\ell(X \wedge \epsilon)]} = \theta_0$ regardless of the value of ϵ, relies explicitly on the form of the cumulative distribution function of a newsboy distribution. As such, this is a distinctive feature of this class of distributions.

Theorem 2 establishes the Bayesian consistency of the posterior distribution $\{\Theta \,|\, \mathcal{F}_N\}$ at the true value θ_0 (DeGroot 2005). This implies that the true value will be learned with certainty as information is accumulated.

5.2. Convergence of myopic policy

Given full knowledge of the true value θ_0, the optimal age-replacement threshold for each component can be computed by minimizing the direct expected discounted cost function, that is,

$$\begin{aligned}
\tau^*(\theta_0) &:= \arg\min_{\tau \geq 0} C(\tau \,|\, \theta_0) \qquad (8) \\
&= \arg\min_{\tau \geq 0} \left\{ c_c \int_0^\tau e^{-\alpha x} f_X(x \,|\, \theta_0)\mathrm{d}x + c_p e^{-\alpha\tau}(1 - F_X(\tau \,|\, \theta_0)) \right\},
\end{aligned}$$

where we use the notation $C(\tau \,|\, \theta_0)$ to denote the direct expected discounted cost function when θ_0 is known. We refer to $\tau^*(\theta_0)$ as the Oracle as this decision requires full knowledge about the unknown parameter and is hence not attainable in practice. The following remark relates the uniqueness and finiteness of $\tau^*(\theta_0)$ to properties of $\ell(x)$.

Remark 1. It has been shown in Fox (1966) that $\tau^*(\theta_0)$ is unique and finite if and only if X has a strictly increasing hazard rate. An increasing hazard rate implies that the component degrades over time so that there is an incentive to perform preventive maintenance. Given full knowledge about θ_0, the hazard rate, denoted with $h(x\,|\,\theta_0)$, is equal to

$$h(x\,|\,\theta_0) = \frac{\frac{d}{dx}F_X(x\,|\,\theta_0)}{1 - F_X(x\,|\,\theta_0)} = \frac{\theta_0 \ell'(x) e^{-\theta_0 \ell(x)}}{e^{-\theta_0 \ell(x)}} = \theta_0 \ell'(x).$$

It is then obvious (recall that $\theta_0 > 0$) that $h(x\,|\,\theta_0)$ is strictly increasing if and only if $\ell'(x)$ is strictly increasing. Hence, $\tau^*(\theta_0)$ is unique and finite if and only if $\ell''(x) > 0$. This statement can thus be verified a-priori without any knowledge about the unknown parameter θ_0.

A policy that is attainable in the absence of knowledge on the true value θ_0 is to, upon installment of the $n-$th component, implement the age-replacement threshold that only minimizes the direct expected discounted costs given the accumulated information. Recall that we used this function also in the dynamic program formulation, where (m, k) captured the accumulated information, see Equation (5). In other words, this myopic policy does not integrate learning with decision making, and solely focuses on exploitation. For $n = 1, 2, \ldots, N$, we denote the optimal age-replacement threshold of this myopic Bayesian policy with $\tau^{mb}(\mathcal{F}_n)$, so that

$$\tau^{mb}(\mathcal{F}_n) := \arg\min_{\tau \geq 0} C_n(\tau\,|\,\mathcal{F}_n).$$

The following result establishes the asymptotic optimality of this myopic Bayesian policy in the sense that the induced decision converges to the Oracle. It relies on the condition that the Oracle is unique and finite, which, as stated before, can be easily verified before any information is accrued, see Remark 1.

Theorem 3. *Suppose $\tau^*(\theta_0)$ is unique and finite. Then the myopic Bayesian policy is asymptotically optimal; that is,*

$$\lim_{N\to\infty} \tau^{mb}(\mathcal{F}_N) = \tau^*(\theta_0).$$

Proof. We first prove that $C_N(\tau\,|\,\mathcal{F}_N)$ converges uniformly to $C(\tau\,|\,\theta_0)$ when $N \to \infty$. We then show that this uniform convergence together with properties of $C_N(\tau\,|\,\mathcal{F}_N)$ implies a stronger notion of convergence, namely epi-convergence, which leads directly to the desired result. We have

$$\begin{aligned}
&\sup_{\tau\geq 0}\Big|C_N(\tau\,|\,\mathcal{F}_N) - C(\tau\,|\,\theta_0)\Big| \\
&= \sup_{\tau\geq 0}\Big|c_c \int_0^\tau e^{-\alpha x} f_X(x\,|\,\mathcal{F}_N)\mathrm{d}x + c_p e^{-\alpha\tau}(1 - F_X(\tau\,|\,\mathcal{F}_N)) \\
&\quad - c_c \int_0^\tau e^{-\alpha x} f_X(x\,|\,\theta_0)\mathrm{d}x - c_p e^{-\alpha\tau}(1 - F_X(\tau\,|\,\theta_0))\Big| \\
&= \sup_{\tau\geq 0}\Big|c_c \int_0^\tau e^{-\alpha x}(f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0))\mathrm{d}x \\
&\quad + c_p e^{-\alpha\tau}\int_\tau^\infty (f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0))\mathrm{d}x\Big|
\end{aligned}$$

$$\begin{aligned}
&\leq \sup_{\tau\geq 0}\Big(c_c \int_0^\tau e^{-\alpha x}\Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x \qquad (9)\\
&\quad + c_p e^{-\alpha\tau}\int_\tau^\infty \Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x\Big) \\
&\leq \sup_{\tau\geq 0}\Big(c_c \int_0^\tau e^{-\alpha x}\Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x\Big) \qquad (10)\\
&\quad + \sup_{\tau\geq 0}\Big(c_p e^{-\alpha\tau}\int_\tau^\infty \Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x\Big) \\
&= c_c \int_0^\infty e^{-\alpha x}\Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x \\
&\quad + c_p \int_0^\infty \Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x\Big). \qquad (11)
\end{aligned}$$

Inequality (9) follows from the triangle inequality and Hölder's inequality, and finally, Inequality (10) is a triangle-like inequality for the supremum operator.

Note that

$$f_X(x\,|\,\mathcal{F}_N) = \int_0^\infty f_X(x\,|\,\theta) p_\Theta(\theta\,|\,\mathcal{F}_N)\mathrm{d}\theta.$$

Since $f_X(x\,|\,\theta)$ is a bounded, continuous function, we have by Theorem 2 and the weak convergence of measures that

$$f_X(x\,|\,\mathcal{F}_N) \to f_X(x\,|\,\theta_0), \quad \text{when } N \to \infty. \qquad (12)$$

Using this weak convergence of measures, we have by Scheffé's Theorem (see e.g., Billingsley 1995, Theorem 16.11) that

$$\lim_{N\to\infty}\int_0^\infty \Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x = 0, \qquad (13)$$

$$\lim_{N\to\infty}\int_0^\infty e^{-\alpha x}\Big|f_X(x\,|\,\mathcal{F}_N) - f_X(x\,|\,\theta_0)\Big|\mathrm{d}x = 0. \qquad (14)$$

Using the established bound in (11) in combination with (13) and (14), we have

$$\lim_{N\to\infty}\sup_{\tau\geq 0}\Big|C_N(\tau\,|\,\mathcal{F}_N) - C(\tau\,|\,\theta_0)\Big| = 0,$$

which establishes the uniform convergence of the direct expected cost functions.

Since $C_N(\tau\,|\,\mathcal{F}_N)$ converges uniformly to $C(\tau\,|\,\theta_0)$, and $C_N(\tau\,|\,\mathcal{F}_N)$ is finite (i.e. $0 < C_N(\tau\,|\,\mathcal{F}_N) \leq c_c$) and continuous for all $\tau \geq 0$ and $N \geq 0$, we have by Proposition 7.15 of Rockafellar and Wets (2009) that $C_N(\tau\,|\,\mathcal{F}_N)$ epi-converges to $C(\tau\,|\,\theta_0)$ when $N \to \infty$.

Note that the sequence $C_N(\tau\,|\,\mathcal{F}_N)$ is eventually level-bounded since we assume that $\tau^*(\theta_0)$ is unique and finite (hence the arg min set will eventually be bounded and nonempty). Then, as $C(\tau\,|\,\theta_0)$ is a proper and left semi-continuous function, we have by Theorem 7.33 of Rockafellar and Wets (2009) that

$$\lim_{N\to\infty}\arg\min_{\tau\geq 0} C_N(\tau\,|\,\mathcal{F}_N) = \arg\min_{\tau\geq 0} C(\tau\,|\,\theta_0),$$

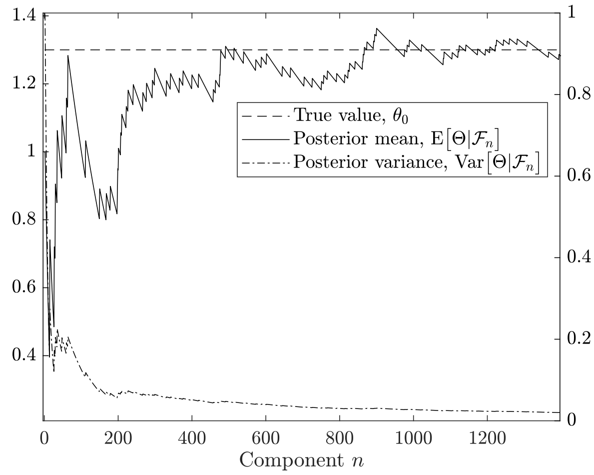
which establishes the result. □

As noted before, computing the optimal policy via the proposed dynamic program is analytically intractable and even numerically a difficult task. As such, decision makers can resort to the myopic Bayesian policy. This policy is not only computationally appealing, but also, as is established in Theorem 3, asymptotically optimal.

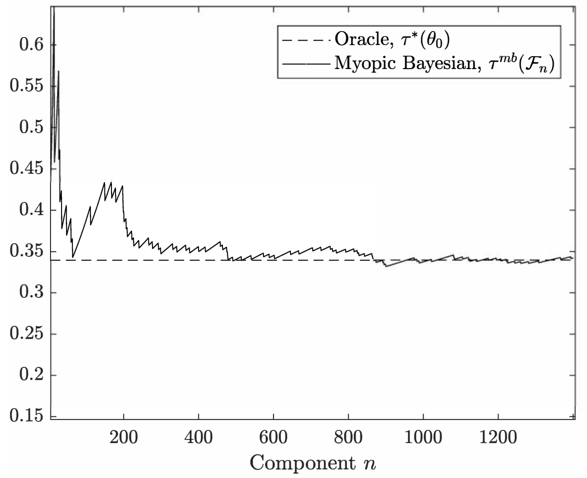
5.3. Illustrative example

Figure 1 provides an illustrative example of the established asymptotic properties for the case when the lifetimes are Weibull distributed random variables with known shape $\beta = 3$ and unknown scale. Since $\beta > 1$, the lifetime distribution has an increasing hazard rate so that the assumption of a unique and finite $\tau^*(\theta_0)$ is justified. The value of $\tau^*(\theta_0)$ is computed using Equation (8) with $\alpha = 0.9$.

Sub-figure (a) shows that the posterior mean converges to the true value θ_0, while the posterior variance converges to 0 at the same time. Sub-figure (b) shows the corresponding sequence of decisions induced by the myopic Bayesian policy, and its convergence to $\tau^*(\theta_0)$. The posterior mean in Figure (a) and the myopic Bayesian policy in Figure (b) appear to be closely linked. The coupling between the evolution of the posterior mean and the myopic Bayesian policy can be explained intuitively. For instance, the jump in the myopic Bayesian policy starting at $n \approx 80$ can be explained as follows. If the posterior mean of Θ decreases based on the accumulated information, the decision maker expects that the posterior predictive lifetime is becoming larger in expectation (see Proposition 1), hence, a higher age-replacement threshold is imposed. The reverse holds true as well, as is nicely illustrated when the posterior mean starts to increase again after $n \approx 200$.



(a) Sequence of posterior mean (left scale) versus the true value θ_0, and sequence of posterior variance (right scale).



(b) Sequence of decisions induced by myopic Bayesian policy versus the decision induced by the Oracle $\tau^*(\theta_0)$.

Figure 1: Illustration of asymptotic properties for the consistency (a) and the decision making (b) when $c_p = 1$, $c_c = 4$, $\ell(x) := x^3$, $\theta_0 = 1.3$ and $\tau^*(\theta_0) \approx 0.34$. The data points in both sub-figures are obtained from the same sample path.

6. Conclusions

We considered a sequence of age-replacement problems with a general lifetime distribution parametrized by a parameter that is unknown a-priori. By adopting a parametric Bayesian framework often used in the inventory research community, we were able to investigate the exploration-exploitation trade-off that naturally arises when age-replacement decisions are integrated with learning from both censored and uncensored observations.

A new stochastic order for this parametric Bayesian framework was established that is particularly useful in maintenance and reliability related problems. For the case of a finite sequence of components, we then analyzed the optimal policy for a finite sequence of components and established structural properties. For the infinite case, we proposed a computationally appealing myopic policy and proved that it almost surely learns the unknown parameter, and that it converges to the optimal decision one would have taken with full knowledge of the unknown parameter.

Two immediate directions for future research are (i) to investigate the rate of convergence of the asymptotic properties due to its practical importance, and (ii) to study how the myopic decision relates to the optimal decision in the finite sequence when the information state in both is the same. With respect to the latter, we expect that the myopic threshold will be lower as it only focuses on exploitation and neglects the exploratory benefits that a higher threshold has.

Acknowledgements

The research of Collin Drent is supported by the Data Science Flagship framework, a cooperation between the Eindhoven University of Technology and Philips. The research of Stella Kapodistria and Onno Boxma is supported by the NWO Gravitation Programme NETWORKS (Grant number 024.002.003).

References

Ahmad, R., Kamaruddin, S.. An overview of time-based and condition-based maintenance in industrial application. Computers & Industrial Engineering 2012;63(1):135–149.

Barlow, R., Hunter, L.. Optimum preventive maintenance policies. Operations Research 1960;8(1):90–100.

Bensoussan, A., Çakanyıldırım, M., Royal, A., Sethi, S.P.. Bayesian and adaptive controls for a newsvendor facing exponential demand. Risk and Decision Analysis 2009;1(4):197–210.

Bertsekas, D.P., Shreve, S.E.. Stochastic Optimal Control: The Discrete Time Case. Academic Press, 1978.

Billingsley, P.. Probability and Measure. 3rd ed. John Wiley & Sons, 1995.

Bisi, A., Dada, M., Tokdar, S.. A censored-data multiperiod inventory problem with newsvendor demand distributions. Manufacturing & Service Operations Management 2011;13(4):525–533.

Braden, D.J., Freimer, M.. Informational dynamics of censored observations. Management Science 1991;37(11):1390–1404.

Chen, L.. Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. Operations Research 2010;58(2):396–413.

Chen, L., Plambeck, E.L.. Dynamic inventory management with learning about the demand distribution and substitution probability. Manufacturing & Service Operations Management 2008;10(2):236–256.

Chen, N., Ye, Z.S., Xiang, Y., Zhang, L.. Condition-based maintenance using the inverse Gaussian degradation model. European Journal of Operational Research 2015;243(1):190–199.

Dayanik, S., Gürler, Ü.. An adaptive Bayesian replacement policy with minimal repair. Operations Research 2002;50(3):552–558.

De Jonge, B., Scarf, P.A.. A review on maintenance optimization. European Journal of Operational Research 2019;Forthcoming.

DeGroot, M.H.. Optimal Statistical Decisions. John Wiley & Sons, 2005.

Ding, X., Puterman, M.L., Bisi, A.. The censored newsvendor and the optimal acquisition of information. Operations Research 2002;50(3):517–527.

Drent, C., Drent, M., Arts, J., Kapodistria, S.. Real-time integrated learning and decision making for deteriorating systems; 2020. Working paper.

Elwany, A.H., Gebraeel, N.Z., Maillart, L.M.. Structured replacement policies for components with complex degradation processes and dedicated sensors. Operations Research 2011;59(3):684–695.

Fox, B.. Age replacement with discounting. Operations Research 1966;14(3):533–537.

Fox, B.. Adaptive age replacement. Journal of Mathematical Analysis and Applications 1967;18(2):365–376.

Kim, M.J., Makis, V.. Joint optimization of sampling and control of partially observable failing systems. Operations Research 2013;61(3):777–790.

Lariviere, M.A., Porteus, E.L.. Stalking information: Bayesian inventory management with unobserved lost sales. Management Science 1999;45(3):346–363.

Lu, X., Song, J.S., Zhu, K.. Analysis of perishable-inventory systems with censored demand data. Operations Research 2008;56(4):1034–1038.

Mersereau, A.J.. Demand estimation from censored observations with inventory record inaccuracy. Manufacturing & Service Operations Management 2015;17(3):335–349.

Rockafellar, R.T., Wets, R.J.B.. Variational Analysis. 2nd ed. Springer Science & Business Media, 2009.

Ross, S.M.. Stochastic Processes. 2nd ed. John Wiley & Sons, 1996.

Shaked, M., Shanthikumar, J.G.. Stochastic Orders. Springer Science & Business Media, 2007.

Van Oosterom, C., Peng, H., van Houtum, G.J.. Maintenance optimization for a Markovian deteriorating system with population heterogeneity. IISE Transactions 2017;49(1):96–109.

Wang, H.. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 2002;139(3):469–489.