LIMITING AVAILABILITY OF A ONE-UNIT SYSTEM BACKED BY A SPARE UNDER REPAIR OR PREVENTIVE MAINTENANCE

BRUNO BIETH, LIANG HONG, AND JYOTIRMOY SARKAR

Abstract. We consider a one-unit system under continuous monitoring, aided by an identical spare unit and serviced by a facility that performs repair on a failed unit or preventive maintenance on a recalled unit making it as good as new. We assume instantaneous commencement of service and installation to operation. We find the distribution of the system up time and down time when life-, recall- and service-times have arbitrary probability density functions. Hence, we obtain the limiting availability of the system. Also, we compute the servicing cost per unit time to determine whether preventive maintenance is preferable over a repair only model.

1. Introduction

We consider a one-unit repairable system that is supported by an identical spare unit and is always under continuous monitoring. In the beginning, one unit is put on operation and the other spare unit remains on cold standby. The operating unit may fail or be recalled for preventive maintenance even though it has not failed. Immediately the spare unit is placed on operation (called instantaneous installation to operation), while immediately the failed/recalled unit undergoes repair/preventive maintenance (PM) at the service facility (this is called instantaneous commencement of service).

We assume that the operating unit functions for a random amount of time, which is either the complete lifetime until failure or the censored lifetime until recall, whichever happens first. We also assume that service (repair or PM) takes a random amount of time, after which the unit is restored to a level equivalent to a new unit (this is called the perfect service policy) and becomes a viable spare. The operating unit may fail or be recalled while the other unit is still being serviced, in which case the system fails.

To measure the performance of maintainable systems, often it suffices to study the limiting availability A, which is the probability that the system will be found functional at a distant future time. Under the assumption of continuous life-, recall-, repair-, and PM times, the limiting availability exists and, in view of the Key Renewal Theorem, it is given by

$$A = \frac{\text{MSUT}}{\text{MSUT} \ + \ \text{MSDT}}, \tag{1.1}$$

where MSDT stands for mean system down time and is the mean duration from the moment the system fails until it is revived through repair, and MSUT stands for mean system up time and is the mean duration from the epoch when the failed system is revived to the next system failure (see Barlow and Proschan [1]).

In the literature, there have been several papers that allow recall and PM. Osaki and Asakura [3] allow life-, recall-, repair- and PM time to be arbitrary, but they do not allow recall of the operating unit while the other unit undergoes repair or PM. Gopalan and D'Souza [2], on the other hand, allow recall of a unit while the other unit is under service, but they restrict repair- and PM times to be exponential. Both papers calculate the mean time until the system goes down using the Laplace transformation technique. Zjilstra [6] allows the recall time to be the larger of the service time of the other unit and a fixed horizon T. They minimize the expected cost per unit time or the expected total discounted cost by choosing T. These papers provide general recipes for the Laplace transform of the survival function (SF) of time to system failure, but they do not provide either the analytic expression or numerical computations for the mean time to system failure, except for the case when all distributions are exponential. Smith and Dekker [5] considers a 1-out-of-n (good) PM model in which recall time T is fixed. They compute the approximate expected up time, expected down time, and expected cost per unit time, and choose n and T to maximize the long term economy.

Our model extends the Gopalan and D'Souza [2] model by allowing arbitrary life-, recall-, repair- and PM times. In particular, our model allows recall at any time, even at the risk of letting the system go down, because it is preferable to carry out a PM rather than a repair after the system is revived by completing service on the other unit. This is justified since the recall distribution is completely at our discretion. Our paper provides analytic expressions and discusses how to determine whether the preventive maintenance is preferred, when the servicing cost per unit time is taken into account.

The remainder of this paper is organized as follows: In Section 2, we present the mathematical formulation of the preventive maintenance model, and highlight the steps involved in deriving the SF of the system up time (SUT) and the system down time (SDT). Hence, in view of (1.1), we obtain an expression for the limiting availability. Section 3 gives details of obtaining the Laplace transform of the SF of the SUT. Section 4 gives an example illustrating the numerical implementation. Section 5 discusses the optimal choice of recall time distribution as well as the preference of a PM model over a replacement only model in terms of cost per unit time.

2. Mathematical Formulation

Let the (complete) lifetime X_l of the operating unit have cumulative distribution function (CDF) F_l, the censored lifetime X_c until recall of the operating unit (henceforth called the recall time) have CDF F_c, the repair time Y_r of a failed unit have CDF G_r and the PM service time Y_p of a recalled unit have CDF G_p. Let the corresponding survival functions (SF) be $\bar{F}_l = 1 - F_l, \bar{F}_c, \bar{G}_r, \bar{G}_p$. Also let these random variables be absolutely continuous with probability density functions (PDF) f_l, f_c, g_r, g_p, respectively. We allow the service time to possibly depend on the immediately preceding operation time of the same unit. All other lifetimes, recall times, repair times and PM times are assumed to be stochastically independent.

For the first operating unit, let the lifetime be X_{l1}, recall time X_{c1}, repair time Y_{r1} and PM time Y_{p1}. We assume that X_{l1} and X_{c1} are independent. That is, the operation on the first unit is terminated as soon as one of two independent causes for termination, failure and recall, takes effect. Thus, we observe either X_{l1} or X_{c1}, depending on whether the first unit fails or is recalled, but we do not observe both. Similarly, we observe either Y_{r1} or Y_{p1}, depending on whether the first unit fails or is recalled, but not both. Letting $\delta_1 = 1$ if the first unit fails and $\delta_1 = 0$ if it is recalled, our observable data consists of (δ_1, X_1, Y_1), where

$$(\delta_1, X_1, Y_1) = \begin{cases} (1, X_{l1}, Y_{rl}) & \text{if } X_{l1} \le X_{c1}, \text{ i.e. if first unit fails} \\ (0, X_{c1}, Y_{pl}) & \text{if } X_{l1} > X_{c1}, \text{ i.e. if first unit is recalled.} \end{cases}$$

Let P_r denote the probability that the operating unit fails before it is recalled. Then

$$P_r := P\{\delta_1 = 1\} = P\{X_l < X_c\} = \int_0^\infty f_l(u)\ \bar{F}_c(u)\ du. \tag{2.1}$$

Note that

$$\begin{cases} X_1 = \min\{X_{l1}, X_{c1}\} = \delta_1\ X_{l1} + (1-\delta_1)\ X_{c1} \\ Y_1 = \delta_1\ Y_{r1} + (1-\delta_1)\ Y_{p1}. \end{cases}$$

We call X_1 the operation time and Y_1 the service time of the first unit. Clearly, X_1 and Y_1 are dependent variables. By the independence of lifetime

X_{l_1} and recall time X_{c_1}, the SF of X_1 is given by

$$\bar{F}(t) = \bar{F}_l(t)\ \bar{F}_c(t), \text{ for all } t \geq 0. \tag{2.2}$$

Hence, the PDF of X_1 is given by

$$f(t) = f_l(t)\ \bar{F}_c(t) + f_c(t)\ \bar{F}_l(t), \text{ for all } t > 0. \tag{2.3}$$

Likewise, we define (δ_2, X_2, Y_2) for the second unit. Henceforth, the units alternately take turns to operate, and yield successive observations $(\delta_3, X_3, Y_3), (\delta_4, X_4, Y_4), \ldots$. The observation vectors corresponding to odd indices refer to successive information from the first unit, and those corresponding to even indices refer to successive information from the second unit. In view of the perfect service policy, we assume that

$$\{(\delta_1, X_1, Y_1), (\delta_2, X_2, Y_2), \ldots\}$$

form a sequence of independent and identically distributed (IID) random vectors. To reinterate, Y_i depends on X_i for all $i \geq 1$.

The system breaks down when the operating unit fails or it is recalled but the service on the other unit is not completed. To define the first system breakdown time, let

$$N = \min\{n \geq 1 : X_{n+1} < Y_n\} \tag{2.4}$$

denote the smallest index n for which the nth service time exceeds the $(n+1)$st operation time. Then $N+1$ is a stopping time with respect to the sequence of observation vectors $\{(\delta_1, X_1, Y_1), (\delta_2, X_2, Y_2), \ldots\}$. We call the epochs $S_1 = X_1, S_2 = S_1 + X_2, \ldots, S_N = S_{N-1} + X_N$ installation times, and classify S_i as type r (type p) depending on whether $\delta_i = 1$ ($\delta_i = 0$), because at epoch S_i a new unit is put on operation while the other unit goes on repair (PM). S_i is of type r with probability P_r and of type p with probability $1 - P_r$, where P_r is defined in (2.1). Note that at epoch $S_N + X_{N+1}$ the system breaks down and it is not an installation time. The next installation time is $S_{N+1} = S_N + Y_N$, and it is of type r or type p depending on whether $\delta_{N+1} = 1$ or 0. S_{N+1} is also called the revival time of a down system. Continuing from S_{N+1} onwards we define other installation times and revival times in a similar manner.

In the sequel we make use of the property that the stochastic behavior of the system after installation times of type r are identical. Similarly, the stochastic behavior of the system after installation times of type p are also identical, though not the same as that after installation times of type r. In other words, the embedded discrete-time stochastic process obtained by looking at the state of the system only at installation times is a two-state Markov chain with state space $\{r,\ p\}$ and transition probabilities given by

$$P_{rr} = P_{pr} = P_r,\ P_{rp} = P_{pp} = 1 - P_r. \tag{2.5}$$

This discrete-time Markov chain has a stationary distribution given by $(P_r, 1-P_r)$ for the states r and p, respectively. In fact, the stationary distribution is attained already at the first installation time X_1.

Returning to the continuous time stochastic process, let T_0 denote the time until the system breaks down starting from $t=0$ when one new unit is put on operation and the other new unit is on cold stand by. Clearly, $T_0 = X_1 + X_2 + \cdots + X_{N+1}$. For $j \geq 1$, let T_{rj} (T_{pj}) denote the additional time until the system breaks down starting from the jth installation and whether the jth installation epoch is of type r (type p). For example, if $\delta_j = 1$ then $X_j + \cdots + X_{N+1}$ is denoted by T_{rj} and if $\delta_j = 0$ then $X_j + \cdots + X_{N+1}$ is denoted by T_{pj}. We have the following relationships:

$$\begin{aligned}\{T_0 > t\} &= \{X_1 > t\} \cup \{X_1 \leq t, \delta_1 = 1, T_{r1} > t - X_1\} \\ &\quad \cup \{X_1 \leq t, \delta_1 = 0, T_{p1} > t - X_1\},\end{aligned} \tag{2.6}$$

$$\begin{aligned}\{T_{r1} > t\} &= \{X_2 > t\} \cup \{Y_{r1} \leq X_2 \leq t, \delta_2 = 1, T_{r2} > t - X_2\} \\ &\quad \cup \{Y_{r1} \leq X_2 \leq t, \delta_2 = 0, T_{p2} > t - X_2\},\end{aligned} \tag{2.7}$$

$$\begin{aligned}\{T_{p1} > t\} &= \{X_2 > t\} \cup \{Y_{p1} \leq X_2 \leq t, \delta_2 = 1, T_{r2} > t - X_2\} \\ &\quad \cup \{Y_{p1} \leq X_2 \leq t, \delta_2 = 0, T_{p2} > t - X_2\}.\end{aligned} \tag{2.8}$$

Note that $T_{r1}, T_{r2}, \ldots$ have the same probability distribution and so do $T_{p1}, T_{p2}, \ldots$. Therefore, we drop the second subscript hereafter. Using (2.6)-(2.8), we determine the SF of T_0, T_r, T_p, and hence their expected values. The details are delegated to Section 3.

Having obtained the SF of T_r and T_p, we can get the SF of SUT and its mean as follows. The duration between successive revival times is called a cycle time. Each cycle either begins in state r with probability P_r, in which case the system remains up for a duration T_r, or the cycle begins in state p with probability $1-P_r$, in which case the system remains up for a duration T_p. Hence, between any two successive revival times the SUT has the same distribution as that of $\delta_1\ T_r + (1-\delta_1)\ T_p$ with SF

$$P\{SUT > t\} = P_r\ P\{T_r > t\} + (1-P_r)\ P\{T_p > t\}, \tag{2.9}$$

and the mean SUT is

$$MSUT = P_r\ E[T_r] + (1-P_r)\ E[T_p]. \tag{2.10}$$

Next, we consider system down time. When the system breaks down with a failure or recall of the operating unit while the service on the other unit is still going on, the system remains in the down state until repair/PM is completed. Thereafter, immediately the serviced unit commences operation and the system enters the up state. For example, the first SDT is $D = Y_N - X_{N+1}$. Recall that by definition of N, we have $Y_N > X_{N+1}$. Also note that the installation time S_N immediately prior to a revival time is of type r if $\delta_N = 1$ (which happens with probability P_r) or of type p if $\delta_N = 0$. Hence, the stochastic behavior of an SDT is the same as that of $\delta_N\ D_r + (1-\delta_N)\ D_p$ where D_r (D_p) denote the SDT if a repair (PM) was going on when the system failed. Their SF are easily obtained by using the independence of $Y_r = Y_{rj}$ ($Y_p = Y_{pj}$) and $X = X_{j+1} = \min\{X_{l,j+1}, X_{c,j+1}\}$ as follows:

$$\begin{aligned} P\{D_r > t\} &= P\{Y_r - X > t | Y_r > X\} \\ &= \int_0^\infty P\{Y_r - u > t | Y_r > u\}\ dF(u) \\ &= \int_0^\infty \frac{\bar{G}_r(t+u)}{\bar{G}_r(u)}\ dF(u), \qquad (2.11) \\ P\{D_p > t\} &= \int_0^\infty \frac{\bar{G}_p(t+u)}{\bar{G}_p(u)}\ dF(u). \end{aligned}$$

Thereafter, we obtain

$$E[D_r] = \int_0^\infty P\{D_r > t\}\ dt,\ E[D_p] = \int_0^\infty P\{D_p > t\}\ dt. \qquad (2.12)$$

Therefore, the SF of the SDT is given by

$$P\{D > t\} = P_r\ P\{D_r > t\} + (1-P_r)\ P\{D_p > t\} \qquad (2.13)$$

$$= \quad P_r \int_0^\infty \frac{\bar{G}_r(t+u)}{\bar{G}_r(u)}\ dF(u) + (1-P_r) \int_0^\infty \frac{\bar{G}_p(t+u)}{\bar{G}_p(u)}\ dF(u),$$

and the mean SDT is given by

$$MSDT \quad = \quad P_r\ E[D_r] + (1-P_r)\ E[D_p]. \qquad (2.14)$$

Finally, the limiting availability is obtained from (1.1), (2.10), and (2.14).

3. Distribution of SUT

We derive the SF $\bar{H}_r(t)$ of T_r and $\bar{H}_p(t)$ of T_p starting from (2.7)-(2.8). The PDF of X_1 is given already in (2.3). Hence,

$$\begin{aligned} P\{\delta_1 = 1 | X_1 = u\} &= P\{X_{l1} \le X_{c1} | X_1 = u\} \qquad (3.1) \\ &= \frac{f_l(u)\ \bar{F}_c(u)}{f_l(u)\ \bar{F}_c(u) + f_c(u)\ \bar{F}_l(u)}. \end{aligned}$$

Therefore, the following integral equations hold:

$$\begin{cases} \bar{H}_r(t) = \bar{F}(t) + \int_0^t G_r(u) \left\{ f_l(u)\ \bar{F}_c(u)\ \bar{H}_r(t-u) \right. \\ \qquad\qquad \left. + f_c(u)\ \bar{F}_l(u)\ \bar{H}_p(t-u) \right\} du \\ \bar{H}_p(t) = \bar{F}(t) + \int_0^t G_p(u) \left\{ f_l(u)\ \bar{F}_c(u)\ \bar{H}_r(t-u) \right. \\ \qquad\qquad \left. + f_c(u)\ \bar{F}_l(u)\ \bar{H}_p(t-u) \right\} du. \end{cases} \qquad (3.2)$$

Taking Laplace transform (and suppressing the argument s), the system of equations (3.2) becomes

$$\begin{cases} \bar{H}_r^* = \bar{F}^* + (G_r f_l \bar{F}_c)^* \ \bar{H}_r^* + (G_r f_c \bar{F}_l)^* \ \bar{H}_p^* \\ \bar{H}_p^* = \bar{F}^* + (G_p f_l \bar{F}_c)^* \ \bar{H}_r^* + (G_p f_c \bar{F}_l)^* \ \bar{H}_p^*. \end{cases}$$

Or equivalently,

$$\begin{bmatrix} 1-(G_r f_l \bar{F}_c)^* & -(G_r f_c \bar{F}_l)^* \\ -(G_p f_l \bar{F}_c)^* & 1-(G_p f_c \bar{F}_l)^* \end{bmatrix} \begin{pmatrix} \bar{H}_r^* \\ \bar{H}_p^* \end{pmatrix} = \bar{F}^* \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{3.3}$$

solving which we get expressions for $\bar{H}_r^*(s)$ and $\bar{H}_p^*(s)$. In particular, evaluating (3.3) at $s = 0$ and solving, we obtain

$$\begin{pmatrix} E[T_r] \\ E[T_p] \end{pmatrix} = E[X_1] \begin{bmatrix} 1-(G_r f_l \bar{F}_c)^*(0) & -(G_r f_c \bar{F}_l)^*(0) \\ -(G_p f_l \bar{F}_c)^*(0) & 1-(G_p f_c \bar{F}_l)^*(0) \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{3.4}$$

Next, combining (2.9) and (3.3), the Laplace transform of the SF of the SUT between successive revival times is given by

$$\bar{H}^*(s) = P_r \ \bar{H}_r^*(s) + (1-P_r) \ \bar{H}_p^*(s) =$$
$$\bar{F}^*(s) \ (P_r, \ 1-P_r) \begin{bmatrix} 1-(G_r f_l \bar{F}_c)^*(s) & -(G_r f_c \bar{F}_l)^*(s) \\ -(G_p f_l \bar{F}_c)^*(s) & 1-(G_p f_c \bar{F}_l)^*(s) \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which may be inverted to obtain the SF of the SUT. Also, combining (2.10) and (3.4), we obtain the MSUT between successive revival times as

$$\text{MSUT} = P_r \ E[T_r] + (1-P_r) \ E[T_p] = \tag{3.5}$$
$$E[X_1] \ (P_r, \ 1-P_r) \begin{bmatrix} 1-(G_r f_l \bar{F}_c)^*(0) & -(G_r f_c \bar{F}_l)^*(0) \\ -(G_p f_l \bar{F}_c)^*(0) & 1-(G_p f_c \bar{F}_l)^*(0) \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Finally, the distribution of T_0, the time until the first system failure, is obtained from the relation $T_0 = X_1 + \delta_1 \ T_r + (1-\delta_1) \ T_p$, hence we also have $E[T_0] = E[X_1] + MSUT$.

How does the availability of the PM model compare with that of the simpler repair only model, in which recall for PM is not allowed (hence, G_p is irrelevant)? Note that the results for this simpler model follow immediately from those of the PM model if we simply let the support of f_c escape to infinity. Then (3.5) becomes

$$\begin{aligned} MSUT_0 &= E[X_l] \ [1-(G_r \ f_l)^*(0)]^{-1} \qquad (3.6) \\ &= E[X_l] \ \left[1-\int_0^\infty G_r(u) \ f_l(u) \, du\right]^{-1} \\ &= E[X_l] \ [1-P\{Y_r \le X_l\}]^{-1} = E[X_l] \ / \ P\{Y_r > X_l\}. \end{aligned}$$

Likewise, (2.14) becomes

$$MSDT_0 = E[D_r] = E[Y_r - X_l | Y_r > X_l]. \tag{3.7}$$

Combining (3.6) and (3.7), the limiting availability, when recall is not allowed, is

$$\begin{aligned} A_0 &= \frac{E[X_l]/P\{Y_r > X_l\}}{E[X_l]/P\{Y_r > X_l\} + E[Y_r - X_l | Y_r > X_l]} \\ &= \frac{E[X_l]}{E[X_l] + P\{Y_r > X_l\}\ E[Y_r - X_l | Y_r > X_l]} \\ &= \frac{E[X_l]}{E[\max\{X_l, Y_r\}]}, \end{aligned} \tag{3.8}$$

which agrees with (1.8) in Sen and Bhattacharjee [4]. In the preventive maintenance model we have the opportunity to attain a limiting availability A, no smaller than A_0 and possibly larger, by choosing the recall time distribution appropriately.

4. Examples

In this Section we evaluate the limiting availability A in the preventive maintenance model when the lifetime and recall time distributions are Weibull and repair time and PM time distributions are exponential. In a similar fashion, we also obtained the numerical results for other arbitrary life-, recall-, repair and PM times such as lognormal and gamma. For the sake of brevity, these results are not presented in this paper.

To clarify the convention used in this paper, we mention that the SF of exponential(μ) is taken to be $e^{-\mu t}$ and that of Weibull(ν, λ) to be $e^{-\lambda t^{\nu}}$. Also ν is called the shape parameter and μ and λ the scale parameters.

Example 4.1. *Suppose that X_l, X_c have Weibull distributions with scale parameters λ_l, λ_c, respectively, but with the same shape parameter ν. Assume Y_r and Y_p have exponential distributions with scale parameters μ_r and μ_p, respectively. Define $\lambda = \lambda_l + \lambda_c$. Note that in this case, $X = min\{X_l, X_c\}$ follows a Weibull distribution with scale λ and shape ν; and that (2.1) yields*

$$P_r = P\{X_l < X_c\} = \int_0^{\infty} \lambda_l\, \nu\, u^{\nu-1}\, e^{-(\lambda_l+\lambda_c)u^{\nu}}\, du = \frac{\lambda_l}{\lambda}. \tag{4.1}$$

Let $\kappa_r = P\{Y_r \le X\} = E[G_r(X)] = E[\bar{F}(Y_r)]$ and $\kappa_p = P\{Y_p \le X\} = E[G_p(X)] = E[\bar{F}(Y_p)]$, which can be evaluated numerically. In particular, in the special case when $\nu = 1$; i.e. when X_{l_1} and X_{c_1} are exponential, we have $\kappa_{r,\nu=1} = \mu_r/(\lambda + \mu_r)$ and $\kappa_{p,\nu=1} = \mu_p/(\lambda + \mu_p)$. Also, in the special *case, $\nu = 2$; that is, when X_l and X_c have Rayleigh distribution, we are able to write closed form expressions for κ_r and κ_p as follows:*

$$\begin{aligned}\kappa_{r,\nu=2} &= \sqrt{\frac{\pi}{\lambda}}\,\mu_r\,e^{\frac{\mu_r^2}{4\lambda}}\left(1-\Phi\left(\frac{\mu_r}{\sqrt{2\lambda}}\right)\right)\\ \kappa_{p,\nu=2} &= \sqrt{\frac{\pi}{\lambda}}\,\mu_p\,e^{\frac{\mu_p^2}{4\lambda}}\left(1-\Phi\left(\frac{\mu_p}{\sqrt{2\lambda}}\right)\right),\end{aligned} \tag{4.2}$$

where $\Phi(t)$ is the CDF of the standard Normal distribution.

In this example, (3.4) becomes

$$\begin{aligned}\begin{pmatrix} E[T_r] \\ E[T_p] \end{pmatrix} &= \frac{\Gamma(1+1/\nu)}{\lambda^{1/\nu}} \begin{bmatrix} 1-\frac{\lambda_l}{\lambda}\kappa_r & -\frac{\lambda_c}{\lambda}\kappa_r \\ -\frac{\lambda_l}{\lambda}\kappa_p & 1-\frac{\lambda_c}{\lambda}\kappa_p \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\\ &= \frac{\Gamma(1+1/\nu)}{\lambda^{1/\nu}} \frac{\begin{pmatrix} 1-\frac{\lambda_c}{\lambda}(\kappa_p-\kappa_r) \\ 1+\frac{\lambda_l}{\lambda}(\kappa_p-\kappa_r) \end{pmatrix}}{1-\frac{\lambda_l}{\lambda}\kappa_r-\frac{\lambda_c}{\lambda}\kappa_p}.\end{aligned} \tag{4.3}$$

Hence, we obtain MSUT from (2.10), (3.4)-(4.1) as

$$MSUT = \frac{\Gamma(1+1/\nu)}{\lambda^{1/\nu}}\left(1-\frac{\lambda_l}{\lambda}\,\kappa_r-\frac{\lambda_c}{\lambda}\,\kappa_p\right)^{-1}. \tag{4.4}$$

By the lack of memory property of the exponential distribution, D_r is exponential(μ_r) and D_p is exponential(μ_p). Hence, MSDT is given by

$$MSDT = \int_0^\infty P\{D>t\}\,dt = \frac{1}{\lambda}\left(\frac{\lambda_l}{\mu_r}+\frac{\lambda_c}{\mu_p}\right); \tag{4.5}$$

and (1.1) simplifies to

$$A = \left[1+\frac{1}{\lambda}\left(\frac{\lambda_l}{\mu_r}+\frac{\lambda_c}{\mu_p}\right)\frac{\lambda^{1/\nu}}{\Gamma(1+1/\nu)}\left(\frac{\lambda_l}{\lambda}(1-\kappa_r)+\frac{\lambda_c}{\lambda}(1-\kappa_p)\right)\right]^{-1}. \tag{4.6}$$

Specializing to the model in which recall for PM is not allowed, (4.6) simplifies to

$$A_0 = \left[1+\frac{(1-\kappa_r)\,\lambda_l^{1/\nu}}{\mu_r\,\Gamma(1+1/\nu)}\right]^{-1}, \tag{4.7}$$

which agrees with (3.8), as $E[X_l] = \lambda_l^{-1/\nu}\Gamma\left(1+\frac{1}{\nu}\right)$ and

$$\begin{aligned}&E[\max\{X_l, Y_r\}]\\ &= E[X_l] + P\{Y_r > X_l\}\; E[Y_r - X_l | Y_r > X_l]\\ &= E[X_l] + \frac{(1-\kappa_r)}{\mu_r}.\end{aligned}$$

Table 1 gives some numerical values.

Table 1: Values of P_r, MSUT, MSDT, A and A_0, in Example 4.1.

		Shape $\nu = 1$				Shape $\nu = 2$				Shape $\nu = 3$			
$(\lambda_l, \lambda_c, \mu_r, \mu_p)$	P_r	MSUT	MSDT	A	A_0	MSUT	MSDT	A	A_0	MSUT	MSDT	A	A_0
(1, 1, 1, 1)	$\frac{1}{2}$	0.7500	1.0000	.4286	.6667	1.1154	1.0000	.5273	.6611	1.3936	1.0000	.5822	.6744
(1, 1, 1, 2)	$\frac{1}{2}$	0.8571	0.7500	.5333	.6667	1.3831	0.7500	.6484	.6611	1.8080	0.7500	.7068	.6744
(1, 1, 2, 2)	$\frac{1}{2}$	1.0000	0.5000	.6667	.8571	1.8200	0.5000	.7845	.8798	2.5732	0.5000	.8373	.8970
(1, 1, 2, 3)	$\frac{1}{2}$	1.1111	0.4167	.7273	.8571	2.1965	0.4167	.8406	.8798	3.2693	0.4167	.8870	.8970
(1, 2, 1, 1)	$\frac{1}{3}$	0.4444	1.0000	.3077	.6667	0.8251	1.0000	.4521	.6611	1.1216	1.0000	.5287	.6744
(1, 2, 1, 2)	$\frac{1}{3}$	0.5128	0.6667	.4348	.6667	1.0682	0.6667	.6157	.6611	1.5588	0.6667	.7004	.6744
(1, 2, 2, 2)	$\frac{1}{3}$	0.5556	0.5000	.5263	.8571	1.2528	0.5000	.7147	.8798	1.9360	0.5000	.7947	.8970
(1, 2, 2, 3)	$\frac{1}{3}$	0.6250	0.3889	.6164	.8571	1.5756	0.3889	.8020	.8798	2.6249	0.3889	.8710	.8970
(2, 1, 1, 1)	$\frac{2}{3}$	0.4444	1.0000	.3077	.4286	0.8251	1.0000	.4521	.5273	1.1216	1.0000	.5287	.5822
(2, 1, 1, 2)	$\frac{2}{3}$	0.4762	0.8333	.3636	.4286	0.9311	0.8333	.5277	.5273	1.3046	0.8333	.6102	.5822
(2, 1, 2, 2)	$\frac{2}{3}$	0.5556	0.5000	.5263	.6667	1.2528	0.5000	.7147	.7845	1.9360	0.5000	.7947	.8373
(2, 1, 2, 3)	$\frac{2}{3}$	0.5882	0.4444	.5696	.6667	1.3957	0.4444	.7585	.7845	2.2284	0.4444	.8337	.8373
(1, 1, 1, 10)	$\frac{1}{2}$	1.2000	0.5500	.6857	.6667	2.0965	0.5500	.7922	.6611	2.7288	0.5500	.8323	.6744
(1, 2, 1, 15)	$\frac{1}{3}$	0.9231	0.3778	.7096	.6667	1.8717	0.3778	.8321	.6611	3.3041	0.3778	.8974	.6744
(1, 1, 2, 20)	$\frac{1}{2}$	1.6923	0.2750	.8602	.8571	3.5401	0.2750	.9279	.8798	5.1189	0.2750	.9490	.8970
(1, 2, 2, 25)	$\frac{1}{3}$	1.2281	0.1933	.8640	.8571	3.5939	0.1933	.9490	.8798	5.7670	0.1933	.9676	.8970

5. Optimal Choice of F_c and Preference of a PM Model via Cost Analysis

Notice that the choice of the recall time distribution F_c is entirely at the discretion of the reliability engineer, all other distributions being dictated by prevailing technology and environmental conditions. What can we say about the optimal choice of F_c that would maximize A in the PM model? We cannot give a comprehensive answer using the tools developed in this paper. The optimal choice of F_c necessitates allowing its support to be a (strict) subset of $(0, \infty)$ or even a discrete set of points. Since our model assumes continuity of F_c from the outset, we delegate the optimal choice problem to a future work. Nonetheless, should we choose F_c to be continuous, we can inquire about the optimal choice of the parameter(s) of F_c. For instance, in Example 4.1 with $\nu = 1$, the optimal value of λ_c may be obtained by maximizing A in (4.6) with respect to $\lambda_c \geq 0$, given the values of λ_l, μ_r, μ_p. For the top twelve values of λ_l, μ_r, μ_p in Table 1 when $\nu = 1$, the optimal choice is $\lambda_c^* = 0$; that is, it is best not to exercise the option to recall. For the last four values of λ_l, μ_r, μ_p, the optimal values of λ_c^* are given in Table 2 below. Note that when the lifetime and repair time distributions remain the same, as the mean time for PM decreases, the optimal recall time distribution becomes stochastically smaller; that is, an earlier recall leads to an increase in limiting availability.

Table 2: The optimal choice of λ_c^*, given λ_l, μ_r, μ_p, and associated values of P_r, MSUT, MSDT and limiting availability A^*, in the special case of Example 4.1 with $\nu = 1$.

λ_l	λ_c^*	μ_r	μ_p	P_r	MSUT	MSDT	A^*	A_0
1	0.9442	1	10	.5144	1.2284	0.3138	.6858	.6667
1	1.7013	1	15	.3702	1.0106	0.4122	.7103	.6667
1	0.7059	2	20	.5862	1.9386	0.3355	.8607	.8571
1	1.2452	2	25	.4454	1.5835	0.2449	.8661	.8571

The PM model is also preferable in situations where the cost M of each PM task is substantially lower than the cost R of each repair, even if A may be slightly smaller than A_0. The (long run) servicing cost per unit time is the ratio of expected total servicing cost within a cycle to the expected length of a cycle. The servicing cost per unit time for the PM model is

$$\frac{E[N]\ \{P_r\ R + (1 - P_r)\ M\}}{\text{MSUT}\ +\ \text{MSDT}}, \tag{5.1}$$

while for the model without PM it is

$$\frac{E[N_0]\ R}{E[N_0]\ E[X_l] + E[Y_r - X_l | Y_r > X_l]}, \tag{5.2}$$

where N and N_0 are the number of units that operate between successive system failures in the two models. Since $N+1$ is a stopping time, Wald's First Identity and (3.5) yield

$$E[N] = (P_r,\ 1-P_r)\ \begin{bmatrix} 1-(G_r f_l \bar{F}_c)^*(0) & -(G_r f_c \bar{F}_l)^*(0) \\ -(G_p f_l \bar{F}_c)^*(0) & 1-(G_p f_c \bar{F}_l)^*(0) \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{5.3}$$

Letting $f_c = 0$ in (5.3), we have

$$E[N_0] = \frac{1}{P\{Y_r > X_l\}}, \tag{5.4}$$

which can be seen also from the fact that N_0 is a geometric random variable with success probability $P\{Y_r > X_l\}$. Substituting (5.3) in (5.1) and (5.4) in (5.2) we evaluate the servicing cost per unit time in the two models. The PM model is preferable if (5.1) is smaller than (5.2).

Thus, a practitioner can decide to recall equipment for preventive maintenance if that would increase the system limiting availability and also decrease the servicing cost per unit time. In case these desirable features are at odds with each other, the practitioner can make a judgment call based on their priorities.

6. Acknowledgment

We sincerely thank the anonymous reviewer whose detailed comments contributed to an improved version of this paper.

References

[1] R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart & Winston, New York, 1975.

[2] M. N. Gopalan and C. A. D'Souza, *Probabilistic analysis of a two-unit system with a single service facility for preventive maintenance and repair*, Operations Research, **23.1** (1975), 173–177.

[3] S. Osaki and T. Asakura, *A two-unit standby redundant system with repair and preventive maintenance*, J. Appl. Prob., **7** (1970), 641–648.

[4] P. K. Sen and M. C. Bhattacharjee, Non-parametric estimators of availability under provisions of spare and repair, *Reliability and Quality Control*, A. P. Basu (ed.), Elsevier, Amsterdam, 1986, 281–296.

[5] M. A. J. Smith and R. Dekker, *Preventive maintenance in a 1 out of n system: The uptime, downtime and costs*, European J. Operational Research, **99** (1997), 565–583.

[6] M. Zjilstra, *Optimization of preventive maintenance for two-unit cold standby redundant systems with a general cost structure*, Zeitschrift für Operations Research, **27** (1982), 55–63.

MSC2000: 90B25, 62N05

Modeling & Simulation, Novartis Pharma AG, Basel, Switzerland
E-mail address: bruno.bieth@novartis.com

Department of Mathematics, Bradley University, Peoria, IL 61625
E-mail address: lhong@bradley.edu

Department of Mathematical Sciences, IUPUI, Indianapolis, IN 46202
E-mail address: jsarkar@math.iupui.edu