Optimal Periodic Inspection Programs for Randomly Failing Equipment

George H. Weiss*

(July 16, 1963)

There have been many analyses made of models for equipment inspection, i.e., where a system may suffer a breakdown, but such an event is only discovered by an inspection. Most analyses assume that the time to failure follows a negative exponential law which implies that only periodic inspection programs need be considered. Another model which has been analyzed by Barlow, Hunter, and Proschan finds the optimal program of inspections when the equipment reliability function is of general form, but a particular loss function is given. In this paper we find the optimal periodic inspection program for systems which do not have negative exponential reliability functions. These programs have the virtue of simplicity even though they may not be optimal in an absolute sense. Besides the periodic inspection programs, we derive results for random inspection programs.

1. Introduction

There have been many papers published on the subject of determining optimum checkout intervals for systems with random failure characteristics, [1–5].[1] The objective of any checkout program is to minimize the down time due to system failure. When the reliability function which describes the system's failure characteristics is negative exponential the optimal checkout policy must be a periodic one. If the system reliability function differs from the exponential, one does not expect such a simple result to obtain. Indeed Barlow, Proschan, and Hunter [3, 4] have devised an algorithm for calculating the optimum checkout schedule for a system with a general reliability function. They minimized the total loss incurred by inspection costs and the cost of uninspected down time to derive an optimum policy. This policy is, however, a difficult one to actually compute since all of the inspection times are a function of the first inspection time. This parameter must be varied in order to find the optimal policy. It has been found that the optimal inspection periods are extremely sensitive to the choice of the first period, hence a good deal of accuracy is needed in the calculations. Other studies of the problem have used the assumption of an exponential reliability function and found the optimal periodic solution [1–5].

It is the purpose of this paper to find the optimal periodic solution for the checkout interval of a system which does not necessarily have an exponential reliability function. Clearly, such a policy will not be optimal in the sense of Barlow, Proschan, and Hunter's results. However the periodic inspection program does have the advantage of simplicity and it is the sort that would most likely be followed in practice. The exact class of policies which will concern us will

be denoted by (Δ, T) policies. The parameter Δ is the inspection interval and T is a time at which the system is to be replaced by a new one if it has passed all of the previous inspections. Three possibilities may be considered: Δ can be held fixed and T varied; T can be held fixed and Δ varied; or both parameters can be varied. In the case of an exponential reliability function or no wearout, (Δ, ∞) policies would be the only logical ones to consider if inspection is considered to be perfect. We may wish to consider this possibility for the general reliability function as well. In addition to an analysis of periodic inspection policies we shall also consider random inspection policies in which the periods between successive inspections are identically distributed random variables.

2. The (Δ, ∞) Policy

Our criterion for determining the best (Δ, T) policy will be the parameter set which maximizes the expected operational readiness of the system. Thus, implicitly the cost function will be measured in lost time; time that is spent in inspection and replacement. The present mathematical model is therefore applicable to computer systems and to other systems which are operated continually.

We assume that system failure is complete, and that the system is replaced rather than repaired when breakdown occurs. The system failure characteristics will be described by a reliability function $R(t)$ (the probability that a system installed at $t = 0$ will survive for a time t or greater). It will be assumed that the first moment

$$\mu = \int_0^\infty R(t)dt \tag{1}$$

is finite. If the inspection and replacement times are assumed to be random variables then one can dis-

*Present Address: Rockefeller Institute, New York, N.Y.

[1] Figures in brackets indicate the literature references at the end of this paper.

tinguish between two types of models. In the first, inspections are scheduled for times $\Delta, 2\Delta, 3\Delta, \ldots$ regardless of the amount of time spent in the last inspection. In the second model the beginning of the next inspection is scheduled Δ units of time after the end of the present one, i.e., inspection is scheduled after a usage period of Δ. If the inspection and replacement times are fixed rather than random the two models are formally identical after a redefinition of Δ. In this paper we will treat the case of a usage period of Δ. Finally, we will assume that the system is turned off during inspection, i.e., that it does not age during an inspection.

The notation in addition to that already defined is:

τ_i = time for a single inspection.
τ_r = time for a single replacement.
θ = probability of not detecting on inspection a system that has failed.

The first case to be studied will be the (Δ, ∞) policy, that is, inspection is continued until a system failure occurs. Let us consider the history of a single system and enumerate the times which will be necessary for our analysis. The nonoperational times will be of three types: An installment time; $n+1$ inspection periods, where n is a random variable; and a period when the system is inoperative but that condition is not yet determined by inspection. There will also be $n+1$ periods of operation, at least n of them being of duration Δ, the remaining one being of duration less than or equal to Δ. For a single system the time history is summarized in figure 1. Let T be the lifetime of a single system measured from its installation to the installation of the next system. Then T is the sum of two parts:

$$T = T_0 + T_N \tag{2}$$

where T_0 is the operational time and T_N is nonoperational time. Barlow and Hunter have shown in [6] that the operational readiness, or proportion of time that the system is operational, is given in the limit of $t \rightarrow \infty$ by

$$P = \frac{E(T_0)}{E(T_0) + E(T_N)}. \tag{3}$$

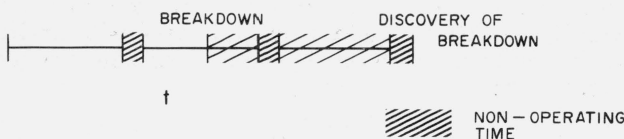where $E(T)$ is the expected value of T. This result is also a simple consequence of renewal theoretic considerations, [7]. Hence in order to evaluate P we must only calculate $E(T_0)$ and $E(T_N)$ for a single system.

The calculation of $E(T_0)$ is a simple one, since it is just the operational time till failure. This is just

$$E(T_0) = \mu \tag{4}$$

where μ has been assumed finite. The calculation of $E(T_N)$ is slightly more complicated. T_N can be written

$$T_N = T_r + T_1 + T_2 + \ldots + T_{n+1} + \zeta \tag{5}$$

where T_r is the replacement, or installation time, T_1, $T_2, \ldots, T_{n+1}$ are the inspection times till just after the failure (n is a random variable) and ζ is the time lost due to faulty inspection. If the expectation of this equation is taken, we find

$$E(T_N) = \tau_r + \tau_i E(n+1) + E(\zeta) \tag{6}$$

where the τ's have already been defined. The random variable n has a distribution

$$Pr\{\text{exactly } n \text{ operational periods}\} = R(n\Delta) - R((n+1)\Delta). \tag{7}$$

Hence $E(n+1)$ is

$$E(n+1) = \sum_{n=0}^{\infty} (n+1) \left[R(n\Delta) - R((n+1)\Delta) \right] \tag{8}$$

$$= \sum_{n=0}^{\infty} R(n\Delta) = G(\Delta)$$

where the finiteness of the first moment insures the convergence of this series. Finally, the residual nonoperational time is composed of the random time between the failure and the first inspection thereafter and a random period due to the effect of imperfect inspection. The probability density of a system failure is $-R'(t)$, which implies that $E(\zeta)$ is given by

$$E(\zeta) = -\sum_{n=0}^{\infty} \int_{n\Delta}^{(n+1)\Delta} ((n+1)\Delta - t)R'(t)dt$$

$$+ (1-\theta) \sum_{m=0}^{\infty} m\theta^m(\tau_i + \Delta)$$

$$= \sum_{n=0}^{\infty} \Delta R(n\Delta) - \mu + \frac{\theta}{1-\theta}(\tau_i + \Delta) \tag{9}$$

$$= \Delta G(\Delta) - \mu + \frac{\theta}{1-\theta}(\tau_i + \Delta).$$

Combining eqs (3), (6), (8), and (9), we find for the operational readiness

$$P(\Delta) = \frac{\mu}{\tau_r + (\tau_i + \Delta)\left(G(\Delta) + \frac{\theta}{1-\theta}\right)}$$

$$= \frac{1}{\lambda_r + \left(\lambda_i + \frac{\Delta}{\mu}\right)\left(G(\Delta) + \frac{\theta}{1-\theta}\right)} \tag{10}$$

FIGURE 1. *Representation of the history of a piece of equipment.*

where we have introduced the dimensionless parameters

$$\lambda_r = \tau_r/\mu, \ \lambda_i = \tau_i/\mu. \tag{11}$$

Occasionally useful bounds on $P(\Delta)$ can be obtained by use of the inequalities

$$\Delta + \mu > \Delta G(\Delta) > \mu \tag{12}$$

This is easily derived by reference to figure 2. There it is seen that $\Delta G(\Delta)$ is the area under the upper set of rectangles, which is greater than the area under the curve for $R(t)$. Furthermore the area under the curve for $R(t)$ is greater than that under the lower set of rectangles. This latter area is, however $\Delta \sum_{n=1}^{\infty} R(n\Delta)$ $= \Delta(G(\Delta)-1)$ from which the upper bound follows. The inequalities of eq (12) thus yield bounds on $P(\Delta)$ as follows

$$\frac{1}{\lambda_r + \left(\lambda_i + \dfrac{\Delta}{\mu}\right)\left(\dfrac{\mu}{\Delta} + \dfrac{1}{1-\theta}\right)} < P(\Delta)$$

$$< \frac{1}{\lambda_r + \left(\lambda_i + \dfrac{\Delta}{\mu}\right)\left(\dfrac{\mu}{\Delta} + \dfrac{\theta}{1-\theta}\right)}. \tag{13}$$

If we write this inequality as $L(\Delta) < P(\Delta) < U(\Delta)$ then it is interesting to note that both $L(\Delta)$ and $U(\Delta)$ have maxima. The maximum of $L(\Delta)$ occurs at

$$\Delta_m = \mu\sqrt{\lambda_i(1-\theta)} \tag{14}$$

and the maximum of $U(\Delta)$ occurs at

$$\Delta_m' = \frac{\Delta_m}{\sqrt{\theta}} \tag{15}$$

Thus the maximum value of $P(\Delta)$ obeys the inequalities

$$\frac{1}{\lambda_r + \left(1 + \sqrt{\dfrac{\lambda_i}{1-\theta}}\right)^2} < P_{max}$$

$$< \frac{1}{\lambda_r + \left(1 + \sqrt{\dfrac{\lambda_i\theta}{1-\theta}}\right)^2}. \tag{16}$$

The optimal value of Δ, call it Δ_0, is the solution to

$$G(\Delta_0) + \frac{\theta}{1-\theta} + (\tau_i + \Delta_0)G'(\Delta_0) = 0. \tag{17}$$

and the value of $P(\Delta_0)$ is

$$P(\Delta_0) = \frac{\mu}{\tau_r - (\tau_i + \Delta_0)^2 G'(\Delta_0)}. \tag{18}$$

When τ_i/μ is small in comparison to 1 we can derive an approximate expression for Δ_0. If τ_i were identi-
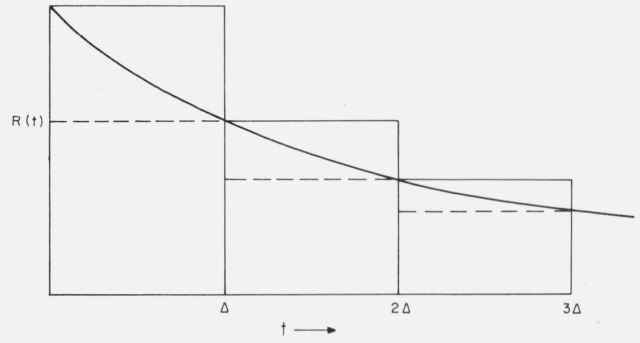

FIGURE 2. *Plot of* R(t) *and upper and lower bounding functions.*

cally equal to zero, that is, if inspection takes no time to carry out, then $\Delta_0 = 0$, or continuous monitoring is the optimal solution. If τ_i/θ is small but not identically equal to zero, it may be supposed that Δ_0/μ will also be a small quantity. Let us then expand $G(\Delta_0)$ and $G'(\Delta_0)$ for small Δ_0 by means of the Euler-Maclauren expansion, [8]

$$G(\Delta) = \sum_{n=0}^{\infty} R(n\Delta) = \frac{\mu}{\Delta} + \frac{1}{2} - \frac{R'(0)}{12}\Delta + 0(\Delta^2)$$

$$G'(\Delta) = -\frac{\mu}{\Delta^2} - \frac{R'(0)}{12} + 0(\Delta). \tag{19}$$

In this approximation the equation for the optimal Δ_0 becomes

$$\frac{\mu}{\Delta_0} + \frac{1}{2} + \frac{\theta}{1-\theta} - \mu\left(\frac{\tau_i}{\Delta_0^2} + \frac{1}{\Delta_0}\right) = 0. \tag{20}$$

We have neglected terms $R'(0)\Delta$ and $R'(0)\tau_i$ which are presumed small in comparison to terms like $1/2$ and $\tau_i\mu/\Delta^2$ in the present calculation. Equation (20) has the solution

$$\frac{\Delta_0}{\mu} = \sqrt{\frac{2(1-\theta)}{1+\theta}\frac{\tau_i}{\mu}} + 0\left(\frac{\tau_i}{\mu}\right) \tag{21}$$

which generalizes a result found earlier for the exponential reliability function [5]. To this approximation the value of P_{max} is

$$P_{max} = \frac{2(1-\theta)}{2\lambda_r(1-\theta) + (\sqrt{\lambda_i(1+\theta)} + \sqrt{2(1-\theta)})^2} \tag{22}$$

In figure 3 we have plotted $P(\Delta/T)$ for the reliability function $R(t) = \exp(-t^2/T^2)$ for Δ/T in the range 0 to 0.20 and $\theta = 0$ and $\theta = 0.25$. It can be observed that the variation of the maximum with Δ/T is negligible in the range that we have considered.

3. Random Inspection Periods

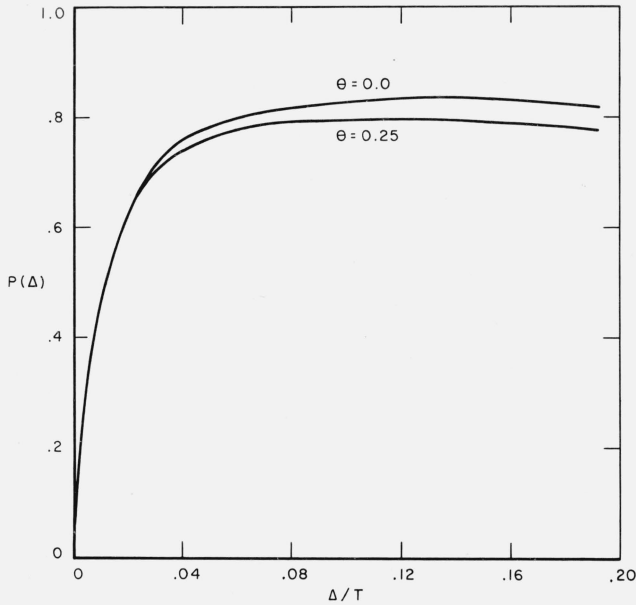Thus far we have considered an idealized case in which inspections are performed as scheduled. It is

FIGURE 3. *Plot of* $P(\Delta)$ *for* $R(t) = exp\ (-t^2/T^2)$ *for* $\theta = 0.25$ *and* $\lambda_i = 0.01$, $\lambda_r = 0.05$.

of some interest to consider the possibility of random examinations, since this, rather than perfect control over inspection is likely to be a more accurate description of actual monitoring situations. We will assume that the inspection following a given one takes place at some time t after the conclusion of the first inspection. The time t is a random variable which will be described by a probability density function $\Psi(t)$ with a finite first moment ν.

Again, in order to calculate operational readiness we may use eq (3) with $E(T_0) = \mu$. The expression for the nonoperating time given in eq (6) is still valid but it requires somewhat more effort to evaluate $E(T_N)$. As before we need to evaluate $E(n+1)$ and $E(\zeta)$. We consider $E(n+1)$ first. For the evaluation of this quantity we define a probability density $w_n(t)$ by

$w_n(t)dt = Pr\{$sum of n periods of operating time is in

$$(t, t+dt)\}. \tag{23}$$

The probability that there are exactly n complete operating periods without a system failure is

$$\Omega_n = \int_0^\infty w_n(t)R(t)dt - \int_0^\infty w_{n+1}(t)R(t)dt \tag{24}$$

since the first integral is the probability that the system is reliable through at least n inspections and the second integral has a similar interpretation. The expectation of $n+1$ is therefore

$$E(n+1) = \sum_{n=0}^\infty (n+1)\left\{\int_0^\infty R(t)\left[w_n(t) - w_{n+1}(t)\right]dt\right\}$$

$$= 1 + \int_0^\infty R(t)w(t)dt \tag{25}$$

where

$$w(t) = \sum_{n=1}^\infty w_n(t) \tag{26}$$

and we have used as a definition

$$w_0(t) = \delta(t). \tag{27}$$

Since the $w_n(t)$ satisfy the recurrence relation

$$w_{n+1}(t) = \int_0^t w_n(\tau)\Psi(t-\tau)d\tau$$

$$w_1(t) = \Psi(t), \tag{28}$$

it is easily verified, by summing both sides of this equation, that $w(t)$ is a solution of the renewal equation

$$w(t) = \Psi(t) + \int_0^t w(\tau)\Psi(t-\tau)d\tau. \tag{29}$$

Notice that $w(t)dt$ is the probability that some inspection occurs after an amount of operating time between t and $t + dt$.

Finally, we must evaluate $E(\zeta)$. Let us assume that there is a system failure after τ hours of operating time. The first component of ζ is the time to the immediately following inspection period. A second component will have to be added in case $\theta \neq 0$. Let the time from the system failure to the immediately following inspection be t (in the terminology of renewal theory this is the forward delay time). In order to calculate the distribution of the forward delay we must take two possibilities into account; either the next inspection is the first one, or else the last inspection ended in the time interval $(x, x+dx)$ with probability $w(x)dx$. These last two events combine to give a result for the forward delay density at time τ, $\rho(\tau, t)$,

$$\rho(\tau, t) = \Psi(\tau + t) + \int_0^\tau w(x)\Psi(\tau + t - x)dx. \tag{30}$$

The first component of $E(\zeta)$ is

$$E_1(\zeta) = -\int_0^\infty d\tau \int_0^\infty dt\, t\rho(\tau, t)R'(\tau) \tag{31}$$

where we have averaged over all possible failure times. The second component of $E(\zeta)$ is the expected down time due to undiscovered failure. The probability that it will take exactly n inspections to discover the failure is $(1-\theta)\theta^{n-1}$ and the expected amount of down time consumed before the discovery is made is

$$E_2(\zeta) = (1-\theta)\sum_{n=1}^\infty \theta^{n-1}(n-1)(\nu + \tau_i) = \frac{\theta(\nu + \tau_i)}{1-\theta} \tag{32}$$

since ν is the expected time between two consecutive inspections. Combining the expressions for $E_1(\zeta)$, $E_2(\zeta)$ and $E(n+1)$ we have the final result

$$P = \frac{\mu}{\tau_r + \mu + \tau_i\left[1 + \int_0^\infty R(t)w(t)dt\right] + \frac{\theta(\nu + \tau_i)}{1 - \theta} - \int_0^\infty d\tau \int_0^\infty dt\, t\rho(\tau, t)R'(\tau)}. \tag{33}$$

In the particular case of a completely random examination policy,

$$\Psi(t) = \sigma e^{-\sigma t} \tag{34}$$

all of the quantities required by eq (33) are readily evaluated. It is well known and easily verified that for the negative exponential distribution one has

$$w(t) = \sigma, \quad \nu = 1/\sigma$$
$$\rho(\tau, t) = \sigma e^{-\sigma t}. \tag{35}$$

Thus we find that

$$-\int_0^\infty R'(\tau)d\tau \int_0^\infty \sigma t e^{-\sigma t}dt = \frac{1}{\sigma} R(\tau)\Big|_0^\infty = \frac{1}{\sigma},$$
$$\int_0^\infty R(t)w(t)dt = \sigma \int_0^\infty R(t)dt = \sigma\mu. \tag{36}$$

The value of P can be given explicitly as

$$P = \frac{\sigma}{\sigma(\lambda_r + 1) + \lambda_i \sigma\left(\frac{1}{1 - \theta} + \sigma\mu\right) + \frac{1}{(1 - \theta)\mu}} \tag{37}$$

where the dimensionless parameters λ_i and λ_r have been defined before. The solution for the maximizing value of σ is

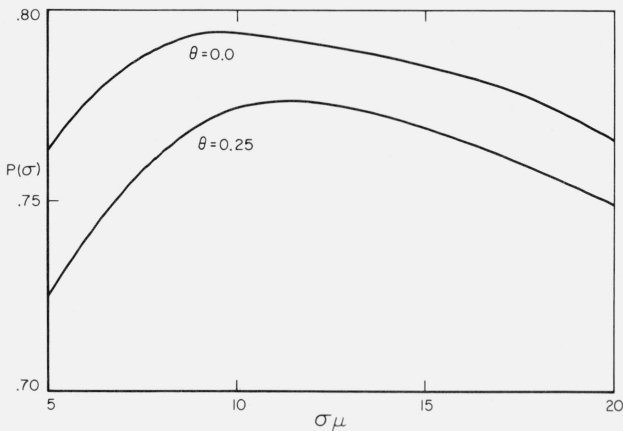$$\sigma_0 = \frac{1}{\mu\sqrt{\lambda_i(1 - \theta)}} \tag{38}$$

FIGURE 4. *Plot of* P *(σ) for random maintenance for $\theta = 0.0, 0.25$ and $\lambda_i = 0.01$, $\lambda_r = 0.05$.*

and the maximum value of P is

$$P_{max} = \left(1 + \lambda_r + \frac{\lambda_i}{1 - \theta} + 2\sqrt{\frac{\lambda_i}{1 - \theta}}\right)^{-1} \tag{39}$$

It is interesting to notice that this result does not depend on $R(t)$ except through μ.

In figure 4 we have plotted P as a function of $\sigma\mu$ for $\lambda_r = 0.05$, $\lambda_i = 0.01$, and $\theta = 0$ and 0.25. It will be observed that the curve is rather flat around the maximum, and the maximum itself does not shift appreciably as θ is changed.

4. The (Δ, NΔ) Case

When the reliability function drops off much more quickly than a negative exponential, it is plausible that one might want to replace the equipment after a specified time rather than continuing to inspect it. Even if the reliability function is a negative exponential, finite scheduled replacement times would be indicated if the quality of inspection is poor, i.e., if θ is close to 1. Quite generally, one might consider a (Δ, NΔ) policy in which replacement takes place after $N\Delta$ units of operating time. In this way the number N becomes an adjustable parameter in the problem, and the policy can be maximized over both Δ and N. This procedure leads to quite complicated expressions. Let us therefore restrict our efforts to a comparison of (Δ, Δ) and (Δ, ∞) policies for particular parameter values.

For the (Δ, Δ) policy we note that

$$E(T_0) = \int_0^\Delta R(x)dx = \mu(\Delta) \tag{40}$$

and

$$E(T_N) = \tau_r + \Delta - E(T_0). \tag{41}$$

These results are to be substituted into eq (3) to obtain the expression for $P(\Delta)$

$$P(\Delta) = \frac{\int_0^\Delta R(x)dx}{\Delta + \tau_r}. \tag{42}$$

In figure 5 we have plotted $P(\Delta)$ as a function of Δ/T for $R(t) = \exp(-t/T)$, for the parameter values $\lambda_r = \tau_r/T = 0.05$ and $\lambda_i = \tau_i/T = 0.01$. When $\theta = 0$ one can do better with a (Δ, ∞) policy than with a (Δ, Δ) policy although when $\theta = 0.9$ this is no longer the case. For other parameter values there may exist an N such that the (Δ, NΔ) policy is better than either the (Δ, Δ) policy or the (Δ, ∞) policy. The value of $P(\Delta)$ for the (Δ, NΔ) policy can be calculated from $E(T_N)$ given by
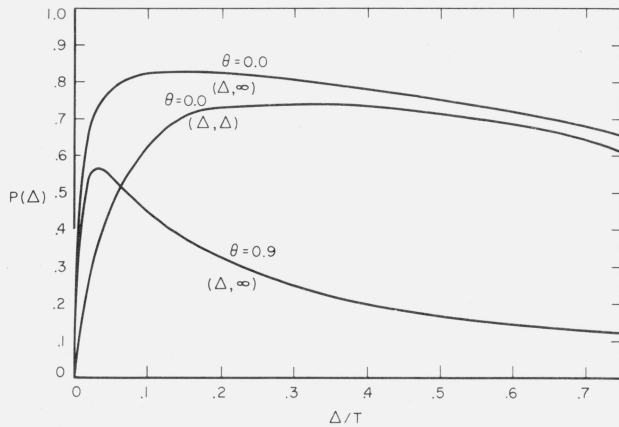
FIGURE 5. *Plot of* $P(\Delta)$ *against* Δ/T *for* $R(x) = \exp(-x/T)$ *for* $\lambda_i = 0.01$, $\lambda_r = 0.05$.

$$E(T_N) = \tau_r - (1-\theta) \sum_{j=0}^{N-2} \sum_{k=1}^{N-j-1} \theta^{k-1}$$

$$\int_{j\Delta}^{(j+1)\Delta} [j\Delta + k(\Delta + \tau_i) - x] R'(x) dx$$

$$- \sum_{j=0}^{N-2} \theta^{N-j-1} \int_{j\Delta}^{(j+1)\Delta} [N\Delta + (N-j-1)\tau_i - x] R'(x) dx$$

$$- \int_{(N-1)\Delta}^{N\Delta} (N\Delta - x) R'(x) dx \qquad (43)$$

and

$$E(T_0) = \int_0^{N\Delta} R(x) dx. \qquad (44)$$

The expression of eq (43) is derived by noting that if the system fails at time x in the interval $[j\Delta, (j+1)\Delta]$ then the time loss due to inspections and to down time is $j\Delta + k(\Delta + \tau_i) - x$ if $k-1$ inspections have been incorrect, accounting for the first set of terms. If all inspections are faulty then the term $1 - \theta$ is to be omitted and the second set of terms results. Finally, if the system fails in $[(N-1)\Delta, N\Delta]$ there are no inspections so that the last term has no θ dependence. The expression for $P(\Delta)$ which results from the use of eqs (43) and (44) is so cumbersome that no information can be extracted from it without considerable computation.

4. References

[1] M. Kamins, Determining checkout intervals for systems subject to random failures, Rand Research Memo RM–2578.
[2] J. Coleman, I. Abrams, Mathematical model for operation readiness, J. Operations Res. **10**, 126 (1962).
[3] R. E. Barlow, L. Hunter, F. Proschan, Optimum monitoring procedures, Boeing scientific Res. Labs. Document, DI–82–0104, 1961.
[4] L. C. Hunter, Optimum checking intervals, in Statistical Theory of Reliability (University of Wisconsin Press) 1963.
[5] G. H. Weiss, A problem in equipment maintenance, Management Science **8**, 266 (1962).
[6] R. E. Barlow, L. C. Hunter, Optimum preventive maintenance policies, J. Operations Res. **8**, 90 (1960).
[7] D. R. Cox, Renewal theory (Methuen, London) 1962.
[8] E. T. Whittaker, G. N. Watson, A course of modern analysis, (Cambridge Press) Fourth edition, 1952.

(Paper 67B4–104)