ADAPTIVE STATISTICAL PROCEDURES IN RELIABILITY AND MAINTENANCE PROBLEMS

BY

J. H. VENTER and J. L. GASTWIRTH

TECHNICAL REPORT NO. 99

OCTOBER 12, 1964

THIS RESEARCH WAS SPONSORED BY THE ARMY RESEARCH OFFICE, OFFICE OF NAVAL RESEARCH, AND AIR FORCE OFFICE OF SCIENTIFIC RESEARCH BY CONTRACT NO. Nonr-225(52) (NR 342-022)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ADAPTIVE STATISTICAL PROCEDURES IN RELIABILITY AND MAINTENANCE PROBLEMS

by

J.L. Gastwirth and J.H. Venter

October 12, 1964

TECHNICAL REPORT NO. 99

PREPARED UNDER CONTRACT Nonr-225(52)

(NR-342-022)

FOR

OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ADAPTIVE STATISTICAL PROCEDURES IN

RELIABILITY AND MAINTENANCE PROBLEMS

by

J.L. Gastwirth and J.H. Venter

1. Introduction.

Consider the following problem: A "system" (or "item", "instrument", etc.) with a "lifetime" that has the distribution function $F(t)$ is to be inspected at times $t_1, t_2, \ldots$. If inspection reveals that the system is inoperative, it is repaired (or replaced); otherwise nothing is done. The general problem is to choose the inspection plan, i.e., the sequence $t_1, t_2, \ldots$, in an optimal way in a suitable sense. Results in this connection can be found in articles [4], [2] where further references can be found. In these studies it is assumed that the distribution function $F(t)$ is known. Although this is usually not the case in practice, the resulting statistical questions have not yet received much attention (see [4], p. 112,113).

In the present paper the case is considered in which the system has an exponential lifetime, i.e.,

$$F(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1-e^{-\lambda t} & \text{if } t \geq 0 , \end{cases} \tag{1.1}$$

with $\lambda (> 0)$ an unknown parameter, and several adaptive or sequential inspections plans are proposed. These plans use information, as it becomes available through inspection, to estimate the unknown parameter

λ and in this manner approach the plan that would be optimum if λ were known. This idea is, of course, not new; it has already been used by Chernoff and others [3], [1] in their papers on sequential design of experiments to which this paper is related.

The plans proposed here are of two general types, viz.,

(i) plans based on maximum likelihood methods,

(ii) plans based on refinements of the Robbins-Monro stochastic approximation method [8].

The asymptotic properties of these two types of plans are generally the same, but from a practical point of view the latter seem preferable in that they are computationally simpler and involve storage of a minimum of past information.

2. Notation and general assumptions.

Unless otherwise stated, we shall assume that there exist known constants $\underline{\lambda}$ and $\overline{\lambda}$ with $0 < \underline{\lambda} < \overline{\lambda} < \infty$ such that

(2.1) $$\lambda \in (\underline{\lambda}, \overline{\lambda}) .$$

Many of the results below can be formulated and proved if λ is restricted only by $0 < \lambda < \infty$, but (2.1) simplifies things considerably and is not unrealistic from a practical point of view in that $\underline{\lambda}$ may be arbitrarily small and $\overline{\lambda}$ arbitrarily large.

We shall also assume that inspections and repairs are instantaneous.

Let $(\Omega, \mathcal{A}, P)$ be a probability measure space. All random variables to be introduced below will be assumed to be defined on $(\Omega, \mathcal{A}, P)$. A generic element of Ω will be denoted by ω and we shall follow the usual practice of exhibiting or omitting the argument ω on random variables according to convenience.

Now consider an inspection plan defined by specifying the inter-inspection times

(2.2) $$T_1 = t_1 , \ T_i = t_i - t_{i-1} , \ i=2,3,\ldots$$

as follows. Let $\{U_n\}$ be an arbitrary sequence of random variables the joint distribution of any finite number of which does not depend on the unknown λ. Take $T_1 = \max\{0,U_1\}$ and define $\{T_n\}$ iteratively by

(2.3) $$T_{n+1} = \max\{0, \ f_n(Y_1,\ldots,Y_n) + U_n\}$$

for $n=1,2,\ldots$. Here, each Y_i, $i=1,2,\ldots$ is a random variable with conditional distribution given $\{Y_1,\ldots,Y_{i-1},T_1,\ldots,T_i\}$ specified by

(2.4) $$Y_i = \begin{cases} 1 & \text{with probability } e^{-\lambda T_i} \\ 0 & \text{with probability } 1-e^{-\lambda T_i} , \end{cases}$$

i.e., $Y_i = 0$ if the i^{th} inspection reveals that the system is inoperative and $Y_i = 1$ otherwise. Also, f_n is a real-valued measurable function of $(Y_1,\ldots,Y_n)$, functionally independent of λ. Intuitively, after n inspections, the next inspection time T_{n+1} depends on the past observations $(Y_1,\ldots,Y_n)$ through f_n while U_n allows for additional randomization.

The class of all these inspection plans will be denoted by $\mathcal{I}$ and a generic element of $\mathcal{I}$ by I.

3. Maximization of information.

We define the average information obtainable from a plan I after n inspections by

(3.1) $$J_n(I,\lambda) = n^{-1}E[\frac{d}{d\lambda} \log L_n(\lambda)]^2$$

where $L_n(\lambda)$ is the likelihood function of λ based on $(Y_1,\ldots,Y_n, T_1,\ldots,T_n)$. Also let,

(3.2) $$J(I,\lambda) = \liminf_{n\to\infty} J_n(I,\lambda) ,$$

and we call $J(I,\lambda)$ the <u>limiting average information</u> obtainable from plan I.

In this section we consider the problem of maximizing $J_n(I,\lambda)$ and $J(I,\lambda)$ by a judicious choice of I. The relevance of this problem to efficient estimation of λ is well known ([5], [6]) and need not be discussed here.

<u>Theorem 3.1</u>: <u>For each</u> n, <u>for all</u> λ <u>and for all</u> I

(3.3) $$J_n(I,\lambda) \leq \lambda^{-1}T_\lambda(2-\lambda T_\lambda)$$

<u>where</u> T_λ <u>is the solution of the equation</u>

(3.4) $$e^{-\lambda T} = 1 - \frac{1}{2}\lambda T .$$

<u>Remark</u>: $T_\lambda = -\lambda^{-1} \log p$ and T_λ is the 100(1-p)-th percentile of the exponential distribution where

(3.5) $$p \doteq .203 .$$

<u>Proof</u>. The conditional probability of observing $Y_1,\ldots,Y_n$ given $(U_1,\ldots,U_n)$ is

(3.6) $$\prod_{i=1}^{n} (e^{-\lambda T_i})^{Y_i}(1-e^{-\lambda T_i})^{1-Y_i} .$$

The distribution of $(U_1,\ldots,U_n)$ being independent of λ, we find

$$\frac{d}{d\lambda} \log L_n(\lambda) = - \sum_{i=1}^{n} T_i Y_i + \sum_{i=1}^{n} (1-Y_i) T_i e^{-\lambda T_i} / (1-e^{-\lambda T_i})$$

$$= \sum_{1}^{n} T_i (Y_i - e^{-\lambda T_i}) / (1-e^{-\lambda T_i})$$

Writing, temporarily,

$$X_i = T_i (Y_i - e^{-\lambda T_i}) / (1-e^{-\lambda T_i}) ,$$

we have, for $j > i$,

$$E(X_i X_j) = E\{X_i T_j (1-e^{-\lambda T_j})^{-1} E[(Y_j - e^{-\lambda T_j}) | Y_1, \ldots, Y_{j-1}, T_1, \ldots, T_j]\}$$

$$= 0$$

because of (2.4). Hence,

$$E[\frac{d}{d\lambda} \log L_n(\lambda)]^2$$

$$= \sum_{1}^{n} E X_i^2$$

$$= \sum_{1}^{n} E\{T_i^2 (1-e^{-\lambda T_i})^{-2} E[(Y_i - e^{-\lambda T_i})^2 | Y_1, \ldots, Y_{i-1}, T_1, \ldots, T_i]\}$$

$$= E \sum_{1}^{n} T_i^2 (1-e^{-\lambda T_i})^{-1} e^{-\lambda T_i}$$

$$\leq n T_\lambda^2 (1-e^{-\lambda T_\lambda})^{-1} e^{-\lambda T_\lambda} ,$$

since the function $T^2(1-e^{-\lambda T})^{-1} e^{-\lambda T}$ is maximized by $T = T_\lambda$. Thus, (3.3) follows.

Equality in (3.3) is attained if and only if $T_i = T_\lambda$ a.s. for each i, i.e., if λ were known, the optimal inspection plan in the sense of maximizing $J_n(I,\lambda)$ for each n and λ would call for periodic inspections with inter-inspection times T_λ. However, within the class $\mathcal{J}$ (i.e., when λ is unknown) there exists no optimal plan.

In order to choose among the plans in $\mathcal{J}$ we have to use a different criterion. Thus, we might require that $\inf_{\lambda\epsilon(\underline{\lambda},\overline{\lambda})} J_n(I,\lambda)$ be maximized for each n, i.e., seek the maximin plan. From (3.3) it follows that such a plan exists in $\mathcal{J}$ and consists of taking $T_i = T_{\overline{\lambda}}$ for each i. This type of criterion does not take into account the information about λ that becomes available as inspection proceeds. Plans with this property can be studied by using the criterion $J(I,\lambda)$.

Definition: An inspection plan I is said to be adaptive (relative to $J(I,\lambda)$) if

$$J(I,\lambda) = \lambda^{-1}T_\lambda(2-\lambda T_\lambda) \quad (3.7)$$

We shall now define and discuss a few adaptive plans.

(i) A Maximum Likelihood Plan.

The following plan is denoted by ML. Let T_1 be an arbitrary number such that

$$T_{\overline{\lambda}} \leq T_1 \leq T_{\underline{\lambda}} .$$

Having defined $T_1,\ldots,T_n$ and observed $Y_1,\ldots,Y_n$ we compute the maximum likelihood estimate λ_n of λ, i.e., if $Y_i=1$ for $i=1,\ldots,n$ let $\mu_n=\underline{\lambda}$ and otherwise let μ_n be the unique solution of the equation

$$\sum_1^n T_i(Y_i-e^{-\lambda T_i})/(1-e^{-\lambda T_i}) = 0 \quad (3.8)$$

in λ; then

$$\lambda_n = \max\{\underline{\lambda},\ \min\{\overline{\lambda},\mu_n\}\} . \quad (3.9)$$

Then we take

(3.10) $$T_{n+1} = T_{\lambda_n} ,$$

n=1,2,... .

<u>Theorem 3.2:</u> <u>The ML plan is adaptive.</u>

<u>Proof.</u> We show firstly that,

(3.11) $$\lambda_n \to \lambda \quad \text{a.s.} \quad \text{as} \quad n \to \infty$$

and for this it suffices to show that

(3.12) $$\mu_n \to \lambda \quad \text{a.s.} \quad \text{as} \quad n \to \infty .$$

Now

$$\begin{aligned} & P\{Y_i=1,\ 1 \le i \le n\} \\ = \ & P\{Y_i=1,\ 1 \le i \le n-1\}P\{Y_n=1 | Y_i=1,\ 1 \le i \le n-1\} \\ \le \ & P\{Y_i=1,\ 1 \le i \le n-1\}e^{-\lambda T_{\underline{\lambda}}} \ . \end{aligned}$$

Iterating backwards we have

$$P\{Y_i=1,\ 1 \le i \le n\} \le e^{-n\lambda T_{\underline{\lambda}}} .$$

Since $P\{Y_i=1,\ 1 \le i < \infty\} \le P\{Y_i=1,\ 1 \le i \le n\}$ for each n, we have

(3.13) $$P\{Y_i=1,\ 1 \le i < \infty\} = 0 .$$

Next let $\mu > \lambda$ and $\delta > 0$ such that $(\mu-\lambda)T_{\bar{\lambda}} > \delta$ and let

(3.14) $$\epsilon = \mu[(\mu-\lambda)T_{\bar{\lambda}} - \delta] .$$

Consider some $\omega\epsilon\Omega$ and suppose that

(3.15) $$\mu < \limsup_n \mu_n(\omega) < \mu + \epsilon .$$

Then there exists a sequence $\{n_k\}$ and an integer k_1 such that for all $k > k_1$

$$\mu \le \mu_{n_k}(\omega) \le \mu + \epsilon$$

and, by (3.13) we may also suppose that for some $i \le n_{k_1}$ $Y_i(\omega) \ne 1$. Hence, from (3.8), for $k > k_1$,

$$0 = \sum_1^{n_k} T_i(Y_i - e^{-\mu_{n_k} T_i})/(1-e^{-\mu_{n_k} T_i})$$

$$\text{(3.16)} \qquad \ge \sum_1^{n_k} T_i Y_i/(1-e^{-(\mu+\epsilon)T_i}) - \sum_1^{n_k} T_i\, e^{-\mu T_i}(1-e^{-\mu T_i})$$

$$= \sum_1^{n_k} T_i e^{-\lambda T_i}/(1-e^{-(\mu+\epsilon)T_i}) - \sum_1^{n_k} T_i e^{-\mu T_i}(1-e^{-\mu T_i}) + \xi_{n_k}$$

where

$$\xi_n = \sum_1^{n} T_i(Y_i - e^{-\lambda T_i})/(1-e^{-(\mu+\epsilon)T_i}).$$

Writing, temporarily,

$$X_i = T_i(Y_i - e^{-\lambda T_i})/(1-e^{-(\mu+\epsilon)T_i}) ,$$

we have

$$E[X_n | X_1, \ldots, X_{n-1}]$$

$$= E\{E[X_n | Y_1, \ldots, Y_{n-1}, T_1, \ldots, T_n] | X_1, \ldots, X_{n-1}\}$$

$$= E\{T_i(1-e^{-(\mu+\epsilon)T_i})^{-1} E[(Y_n - e^{-\lambda T_n}) | Y_1, \ldots, Y_{n-1}, T_1, \ldots, T_n] | X_1, \ldots, X_n\}$$

$$= 0 \text{ a.s.}$$

because of (2.4). Also,

$$\begin{aligned}\mathrm{Var}(X_n) &= E\{T_n^2(1-e^{-(\mu+\epsilon)T_n})^{-2}E[(Y_n-e^{-\lambda T_n})^2|Y_1,\ldots,Y_{n-1},T_1,\ldots,T_n]\} \\ &= E\{T_n^2(1-e^{-(\mu+\epsilon)T_n})^{-2}e^{-\lambda T_n}(1-e^{-\lambda T_n})\} \\ &\leq 2\lambda^{-2}(1-e^{-(\mu+\epsilon)T_{\bar{\lambda}}})^{-2} \\ &< \infty .\end{aligned}$$

Hence, by Theorem E, p. 387 of [7],

$$n^{-1}\xi_n \to 0 \quad \text{a.s.} \quad \text{as} \quad n\to\infty .$$

In particular, we may suppose that for our ω,

$$n_k^{-1}\xi_{n_k}(\omega) \to 0 \quad \text{as} \quad k \to\infty. \tag{3.17}$$

The first two terms on the right in (3.16) are

$$\sum_1^{n_k} T_i e^{-\lambda T_i}(1-e^{-\mu T_i})^{-1}[(1-e^{-\mu T_i})(1-e^{-(\mu+\epsilon)T_i})^{-1}-e^{-(\mu-\lambda)T_i}]$$

and using the inequalities

$$\begin{aligned}(1-e^{-\mu T_i})(1-e^{-(\mu+\epsilon)T_i})^{-1} &\geq (1+\epsilon/\mu)^{-1} \\ e^{-(\mu-\lambda)T_i} &\leq (1+(\mu-\lambda)T_i)^{-1}\end{aligned}$$

this expression is greater than

$$\begin{aligned}&\sum_1^{n_k} e^{-\lambda T_i}\, T_i[(1+\epsilon/\mu)^{-1} - (1+(\mu-\lambda)T_i)^{-1}] \\ &\geq \delta(\mu-\lambda)^{-1}\sum_1^{n_k} e^{-\lambda T_i} \\ &\geq n_k\delta(\mu-\lambda)^{-1}e^{-\lambda T_{\underline{\lambda}}} ,\end{aligned} \tag{3.18}$$

where we have used (3.14) and the fact that $T_{\bar{\lambda}} \leq T_i \leq T_{\underline{\lambda}}$. Therefore, dividing (3.16) by n_k and letting $k \to\infty$ while using (3.17) and (3.18), we get

$$0 = \liminf_{k\to\infty} n_k^{-1} \sum_1^{n_k} T_i(Y_i - e^{-\mu_{n_k} T_i})/(1-e^{-\mu_{n_k} T_i})$$

$$\geq \delta(\mu-\lambda)^{-1} e^{-\lambda T_{\underline{\lambda}}} > 0$$

which is a contradiction. Hence

$$P\{\mu < \limsup_{n\to\infty} \mu_n < \mu+\epsilon\} = 0 \ .$$

It is also readily shown that $P\{\limsup_{n\to\infty} \mu_n = \infty\} = 0$ (in fact $\mu_n \leq T_{\overline{\lambda}}^{-1} e^{\lambda T_{\underline{\lambda}}}$ a.s.).

Thus it follows that

$$P\{\lambda < \limsup_{n\to\infty} \mu_n\} = 0 \ .$$

By a similar argument we show that

$$P\{\liminf_{n\to\infty} \mu_n < \lambda\} = 0$$

and these two results yield (3.12).

Now, since T_λ is a continuous function of λ, it follows from (3.11) and (3.10) that $T_n \to T_\lambda$ a.s. as $n \to \infty$. Also,

$$J_n(ML,\lambda) = E\, n^{-1} \sum_1^n T_i^2 e^{-\lambda T_i}(1-e^{-\lambda T_i})^{-1}$$

and hence it follows from the dominated convergence theorem that

$$J(ML,\lambda) = T_\lambda^2 e^{-\lambda T_\lambda}(1-e^{-\lambda T_\lambda})^{-1} = \lambda^{-1} T_\lambda(2-\lambda T_\lambda) \ ,$$

concluding the proof of the theorem. It is also possible to modify Wald's proof of the consistency of the maximum likelihood estimator [11] to show that $\lambda_n \to \lambda$ a.s..

(ii) <u>A Stochastic Approximation Plan</u>.

The following plan (denoted by SA) is based on the Robbins-Monro stochastic approximation method [8]; it exploits the fact that T_λ

corresponds to the $100(1-p) \doteq 79.7$-th percentile of the exponential distribution, independent of λ. The plan is defined as follows:

Choose T_1 arbitrary in $[T_{\bar{\lambda}}, T_{\underline{\lambda}}]$ and after defining $T_1, \dots, T_n$ let

$$\lambda_n = - T_n^{-1} \log p ,$$

(3.21) $$A_n = \lambda_n^{-1} p^{-1} = - T_n (p \log p)^{-1} ,$$

and $$T_{n+1} = \max\{T_{\bar{\lambda}}, \min\{T_{\underline{\lambda}}, T_n + n^{-1} A_n (Y_n - p)\}\} ,$$

$$n = 1, 2, \dots .$$

For the rationale behind this plan we refer the reader to [8] and [10].

Theorem 3.3: The SA plan is adaptive.

Proof. It suffices to show that

(3.22) $$T_n \to T_\lambda \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

Writing

$$W_n = n^{-1} A_n (Y_n - e^{-\lambda T_n}) ,$$

it follows from Theorem D, p. 387 of [7] that

(3.23) $$\sum_1^n W_k \quad \text{converges a.s. as} \quad n \to \infty .$$

Now, suppose that for some $\omega \epsilon \Omega$ for which (3.23) holds we can find a sequence $\{n_k\}$ such that

(3.24) $$T_{n_k}(\omega) \to T_\lambda \quad \text{as} \quad k \to \infty.$$

Then we shall show that

(3.25) $$T_n(\omega) \to T_\lambda \quad \text{as} \quad n \to \infty.$$

Let $\epsilon (> 0)$ be small enough so that

$$(T_\lambda-\epsilon,\ T_\lambda+\epsilon) \subset (T_{\overline{\lambda}},\ T_{\underline{\lambda}}) .$$

There exists an integer k_ϵ such that

$$|T_{n_k}(\omega)-T_\lambda| < \epsilon/2 \quad \text{for all} \quad k > k_\epsilon ,$$

$$|\sum_m^{m+j} W_i| < \epsilon/4 \qquad \text{for all} \quad m > n_{k_\epsilon}, \text{ and all } j > 0 ,$$

and

$$(n_k+\ell)^{-1}A_{n_k+\ell} < \epsilon/4 \quad \text{for all} \quad k > k_\epsilon \quad \text{and}$$

(3.26) $$\text{all} \quad \ell > 0 .$$

Let $k > k_\epsilon$ be fixed and consider the behavior of T_{n_k+m}, $m=0,1,2,\ldots$. Suppose $T_{n_k}-T_\lambda > 0$ (a similar argument will hold otherwise). From (3.21)

$$T_{n_k+1} - T_\lambda = T_{n_k} - T_\lambda + n_k^{-1}A_{n_k}(e^{-\lambda T_{n_k}} - e^{-\lambda T_\lambda}) + W_{n_k}$$

$$< T_{n_k} - T_\lambda + W_{n_k} < \epsilon .$$

If $T_{n_k+1} - T_\lambda > 0$, then repeating,

$$T_{n_k+2} - T_\lambda \leq T_{n_k} - T_\lambda + W_{n_k} + W_{n_k+1} \leq \epsilon .$$

Repeating this argument we find that if

$$T_{n_k+m} - T_\lambda > 0 \quad \text{for} \quad m=0,1,\ldots,\ell-1$$

then

$$|T_{n_k+m} - T_\lambda| < \epsilon \quad \text{for} \quad m=0,1,\ldots,\ell-1 .$$

Suppose that $T_{n_k+\ell} - T_\lambda \leq 0$. Then

$$
\begin{aligned}
0 &\geq T_{n_k+\ell} - T_\lambda \\
&= T_{n_k+\ell-1} - T_\lambda + (n_k+\ell-1)^{-1}A_{n_k+\ell-1}(e^{-\lambda T_{n_k+\ell-1}} - e^{-\lambda T_\lambda}) + W_{n_k+\ell-1} \\
&> -(n_k+\ell-1)^{-1}A_{n_k+\ell-1} + W_{n_k+\ell-1} \\
&\geq -\epsilon/2 .
\end{aligned}
$$

Now we apply the same argument but starting with $T_{n_k+\ell}$ instead of T_{n_k}. It follows that, for all m,

$$|T_{n_k+m} - T_\lambda| < \epsilon ,$$

and (3.25) follows.

To conclude the proof of the theorem we shall now show that for almost all ω, T_λ is a limit point of the sequence $\{T_n(\omega)\}$. Fix ω and let $\{T_{n_r}\}$ be a subsequence converging to T_o, say, with $T_o \in [T_{\bar{\lambda}}, T_{\underline{\lambda}}]$. We shall suppose that $T_o > T_\lambda$, the other case being similar.

Let $\epsilon (> 0)$ be such that $T_\lambda + \epsilon < T_o - \epsilon$. There exists r_ϵ such that

$$|T_{n_r} - T_o| < \epsilon/2 \quad \text{for} \quad r > r_\epsilon ,$$

$$\left|\sum_m^{m+j} W_i\right| < \epsilon/4 \quad \text{for} \quad m > n_{r_\epsilon}, \ j > 0$$

and

$$(n_r+\ell)^{-1}A_{n_r+\ell} < \epsilon/4 \quad \text{for} \quad r > r_\epsilon, \ \ell > 0 . \tag{3.27}$$

Let $c = e^{-\lambda T_\lambda} - e^{-\lambda(T_\lambda+\epsilon)}$. Then, from (3.21), for $r > r_\epsilon$

$$T_{n_r+1} - T_\lambda \leq T_{n_r} - T_\lambda - cn_r^{-1}A_{n_r} + W_{n_r} .$$

If $T_{n_r+1} - T_\lambda \geq \epsilon$, then we repeat to obtain

$$T_{n_r+2} - T_\lambda \leq T_{n_r} - T_\lambda - c[n_r^{-1}A_{n_r} + (n_r+1)^{-1}A_{n_r+1}] + W_{n_r} + W_{n_r+1} .$$

If again $T_{n_r+2} - T_\lambda \geq \epsilon$, we repeat as before; since $\sum_m (n_r+m)^{-1}A_{n_r+m} = \infty$, it follows that there exists an integer ℓ such that $T_{n_r+\ell} - T_\lambda \leq \epsilon$ while $T_{n_r+\ell-1} - T_\lambda \geq \epsilon$. Then, from (3.21)

$$T_{n_r+\ell} - T_\lambda \geq - (n_r+\ell-1)^{-1}A_{n_r+\ell-1} + W_{n_r+\ell-1} \geq -\epsilon ,$$

i.e.,

$$|T_{n_r+\ell} - T_\lambda| \leq \epsilon .$$

Hence, for each ϵ there exist k such that $|T_k - T_\lambda| \leq \epsilon$, i.e., $\{T_n\}$ has a limiting point at T_λ, and the theorem follows.

As far as the criterion $J(I,\lambda)$ is concerned, the ML and SA plans are equivalent; however, the plan leading to a sequence $\{T_n\}$ which converges fastest to T_λ seems preferable in that this plan would generally lead to the largest average information $J_n(I,\lambda)$ for finite n. One possible way to judge the rate of convergence of $\{T_n\}$ is to consider the variance of the asymptotic distribution of $\sqrt{n}\,(T_n - T_\lambda)$. The asymptotic distributions of $\sqrt{n}\,(T_n - T_\lambda)$ for both plans considered [ML and SA] are given in the next theorem.

<u>Theorem 3.4</u>: <u>For both the ML plan and the SA plan, we have</u>

$$\sqrt{n}\,(T_n - T_\lambda) \xrightarrow{\mathcal{L}} N(0, (1-p)p^{-1}\lambda^{-2}) \tag{3.28}$$

$$\sqrt{n}\,(\lambda_n - \lambda) \xrightarrow{\mathcal{L}} N(0, (1-p)p^{-1}T_\lambda^{-2}) , \tag{3.29}$$

<u>where</u> $N(\mu, \sigma^2)$ <u>denotes the normal probability law with mean</u> μ <u>and variance</u> σ^2.

Proof. First consider the ML plan. Substituting μ_n for λ in (3.9) we have

$$(3.30) \qquad \sum_1^n T_i(e^{-\lambda T_i} - e^{-\mu_n T_i})/(1-e^{-\mu_n T_i}) + \xi_n = 0$$

where

$$\xi_n = \sum_1^n T_i(Y_i - e^{-\lambda T_i})/(1-e^{-\mu_n T_i}) .$$

By a central limit theorem for non-independent random variables such as Theorem C, p. 377 of [7], or Lemma 6, p. 377 of [9] or Lemma 4, p. 238 of [10], we have

$$(3.31) \qquad n^{-1/2}\xi_n \xrightarrow{\mathcal{L}} N(0, T_\lambda^2 e^{-\lambda T_\lambda}(1-e^{-\lambda T_\lambda})^{-1}) \equiv N(0, p(1-p)^{-1}T_\lambda^2) .$$

Since $1-e^{-x} = x[1+g(x)]$ where $g(x) \to 0$ as $x \to 0$, (3.30) can be written

$$(3.32) \qquad \sqrt{n}(\mu_n-\lambda)n^{-1}\sum_1^n T_i^2 e^{-\lambda T_i}(1-e^{-\mu_n T_i})^{-1}[1+g((\mu_n-\lambda)T_i)] = -n^{-1/2}\xi_n .$$

Since

$$n^{-1}\sum_1^n T_i^2 e^{-\lambda T_i}(1-e^{-\mu_n T_i})^{-1}[1+g((\mu_n-\lambda)T_i)] \to$$
$$T_\lambda^2 e^{-\lambda T_\lambda}(1-e^{-\lambda T_\lambda})^{-1} \equiv T_\lambda^2 p(1-p)^{-1} \quad \text{a.s.} \quad \text{as } n \to \infty ,$$

(3.30) and (3.29) yield

$$\sqrt{n}(\mu_n-\lambda) \xrightarrow{\mathcal{L}} N(0, (1-p)p^{-1}T_\lambda^{-2}) .$$

Writing $\sqrt{n}(\lambda_n-\lambda) = \sqrt{n}(\mu_n-\lambda) + \epsilon_{1n}$, it follows from (3.10) and (3.12) that $\epsilon_{1n} \to 0$ a.s. (in fact $\epsilon_{1n} = 0$ for n large enough). Hence $\sqrt{n}(\lambda_n-\lambda)$ and $\sqrt{n}(\mu_n-\lambda)$ have the same asymptotic distributions.

Since

$$\frac{dT_\lambda}{d\lambda} = \lambda^{-2} \log p$$

and

$$T_n = T_{\lambda_{n-1}} ,$$

we obtain also

$$\sqrt{n}(T_n - T_\lambda) \xrightarrow{\mathcal{L}} N(0, (1-p)p^{-1}T_\lambda^{-2}\lambda^{-4}(\log p)^2)$$

which is equivalent to (3.28).

Now consider the SA plan. Since $T_n \to T_\lambda$ and $T_{\bar{\lambda}} < T_\lambda < T_{\underline{\lambda}}$ there exists a random variable N with $P(N < \infty) = 1$ such that, for all $n > N$,

$$T_{\bar{\lambda}} < T_n + n^{-1}A_n(Y_n - p) < T_{\underline{\lambda}}$$

and hence, by (3.21), for $n > N$,

$$\begin{aligned} T_{n+1} &= T_n + n^{-1}A_n(Y_n-p) \\ &= T_n + n^{-1}A_n(e^{-\lambda T_n} - e^{-\lambda T_\lambda}) + W_n \qquad (3.32) \\ &= T_n - n^{-1}(1+\epsilon_{2n})(T_n - T_\lambda) + W_n \end{aligned}$$

where $\epsilon_{2n} \to 0$ a.s. as $n \to \infty$.

Multiplying (3.32) by $(n+1)^\gamma$, rearranging terms and writing temporarily

$$X_n = n^\gamma (T_n - T_\lambda) ,$$

we have

$$X_{n+1} = X_n - n^{-1}(1-\gamma+\epsilon_{3n})X_n + (n+1)^\gamma W_n \qquad (3.33)$$

where $\epsilon_{3n} \to 0$ a.s..

If $\gamma < 1/2$, then the argument we used to show that $T_n - T_\lambda \to 0$ with T_n defined by (3.21) can be applied to the sequence $\{X_n\}$ as given by (3.33) to show that $X_n \to 0$ a.s. as $n \to \infty$; hence

$$(3.34) \qquad T_n - T_\lambda = o(n^{-\gamma}) \text{ a.s., with } \gamma < 1/2, \text{ as } n \to \infty .$$

Again, from (3.32)

$$(3.35) \qquad n(T_{n+1} - T_\lambda) = (n-1)(T_n - T_\lambda) + \epsilon_{4n} + nW_n$$

where

$$\epsilon_{4n} = -\epsilon_{2n}(T_n - T_\lambda) = O((T_n - T_\lambda)^2),$$

as $n \to \infty$. By (3.34)

$$(3.36) \qquad |\epsilon_{4n}| = o(n^{-2\gamma}) \text{ a.s. as } n \to \infty.$$

Iterating (3.35) back to $n = N+1$ and dividing by $\sqrt{n}$, we get

$$(3.37) \qquad \sqrt{n}(T_{n+1} - T_\lambda) = n^{-1/2} N(T_N - T_\lambda) + n^{-1/2} \sum_{N+1}^{n} \epsilon_{4k} + n^{-1/2} \sum_{N+1}^{n} k\, W_k .$$

Now

$$n^{-1/2} N(T_N - T_\lambda) \to 0 \text{ a.s. as } n \to \infty,$$

and, by (3.36)

$$n^{-1/2} \sum_{N+1}^{n} \epsilon_{4k} \to 0 \text{ a.s. as } n \to \infty .$$

Also,

$$n^{-1/2} \sum_{N+1}^{n} k\, W_k = n^{-1/2} \sum_{1}^{n} k\, W_k - n^{-1/2} \sum_{1}^{N} k\, W_k$$

and here

$$n^{-1/2} \sum_{1}^{N} k\, W_k \to 0 \text{ a.s. as } n \to \infty.$$

While by the central limit theorem from which (3.31) follows, we have

$$n^{-1/2} \sum_{1}^{n} k\, W_k \xrightarrow{\mathcal{L}} N(0,(1-p)p^{-1}\lambda^{-2}) .$$

These results together with (3.37) yield (3.28) and conclude the proof of the theorem.

Since the two plans have the same asymptotic distribution we are unable to decide between them on this basis. From a practical point of view the SA plan seems preferable. It is not only computationally simpler but on each step only the values of T_n and Y_n are required in order to calculate T_{n+1}, whereas the entire past set of observations is required in the case of the ML plan.

An important question at this point is whether one could construct an adaptive plan with a smaller asymptotic variance for $\sqrt{n}(T_n - T_\lambda)$ than that of the plans discussed. We state the following theorem ([6]) without proof.

Theorem 3.5: For any plan I such that there exists a sequence of positive numbers $\{\sigma_n(\lambda)\}$ such that $\sigma_n(\lambda) \to 0$ as $n \to \infty$ and

$$\sigma_n(\lambda)^{-1}(T_n - T_\lambda) \xrightarrow{\mathcal{L}} N(0,1)$$

(which implies that I is adaptive), we have

$$\liminf_{n\to\infty} n\sigma_n(\lambda)^2 \geq (1-p)p^{-1}\lambda^{-2}$$

for all $\lambda \epsilon (\underline{\lambda}, \overline{\lambda})$, except possibly for λ in a set of 0 Lebesgue measure.

This theorem states that it is essentially not possible to improve on the ML and SA plans if the criterion used is the asymptotic variance of $\sqrt{n}\,(T_n - T_\lambda)$.

4. Minimization of cost.

We shall now suppose that maintenance of the system involves the following costs: Each inspection costs c_1 units; repair of the system if it is inoperative costs c_2 units; and while the system is inoperative a cost of c_3 units per unit time is incurred. In this section we shall study the problem of choosing the inspection plan so as to minimize the long run expected cost per unit time.

More precisely, consider a plan I giving rise to inter-inspection times $\{T_i\}$ with associated inspection times $t_i = \sum_{j \leq i} T_j$. Over the interval $[0,\tau]$ the cost associated with plan I is

$$c_1 N_\tau + c_2 \sum_{i=1}^{N_\tau} (1-Y_i) + c_3 \left(\sum_{i=1}^{N_\tau} V_i + W_\tau \right) , \tag{4.1}$$

where N_τ is defined by

$$t_{N_\tau} \leq \tau < t_{N_\tau+1} , \tag{4.2}$$

and where V_i denotes the time the system has been inoperative over the interval $[t_{i-1}, t_i]$ and W_τ the time it has been inoperative during $[t_{N_\tau}, \tau]$. Hence, the expected cost per unit time over the interval $[0,\tau]$ associated with plan I is

$$C(I,\lambda,\tau) = \tau^{-1} E \sum_{i=1}^{N_\tau} [c_1 + c_2(1-Y_i) + c_3 V_i] + \tau^{-1} c_3 E W_\tau . \tag{4.3}$$

We also write

$$C(I,\lambda) = \limsup_{n \to \infty} C(I,\lambda,\tau) . \tag{4.4}$$

Within the class $\mathcal{I}$ no inspection plan exists having the property that $C(I,\lambda,\tau)$ is minimized for each τ uniformly in λ. We shall, therefore, use $C(I,\lambda)$ as a criterion for evaluating inspection plans. This criterion is closely related to one suggested by Flehinger [4].

Theorem 4.1: *For each* λ *and for each* I ,

$$C(I,\lambda) \geq c_3 + T_\lambda^{*-1}[c_1+d_\lambda(1-e^{-\lambda T_\lambda^*})] \tag{4.5}$$

where

$$d_\lambda = c_2 - c_3\lambda^{-1} \tag{4.6}$$

and T_λ^* *is the unique solution of the equation*

$$e^{-\lambda T}(1+\lambda T) = 1 + c_1 d_\lambda^{-1} ; \tag{4.7}$$

$$\text{if } \lambda < c_3/(c_1+c_2) \text{ then } 0 < T_\lambda^* < \infty ,$$

$$\text{if } \lambda \geq c_3/(c_1+c_2) \text{ then we take } T_\lambda^* = \infty$$

and in this case the second term in (4.5) should be interpreted as 0.

Proof. Consider a fixed τ and define variables X_i by

$$X_i = \begin{cases} 1 & \text{if } \sum_{j=1}^{i} T_j \leq \tau \\ 0 & \text{otherwise .} \end{cases} \tag{4.8}$$

Then (4.3) can be written

$$\begin{aligned} C(I,\lambda,\tau) &= \tau^{-1}E\sum_1^\infty X_i[c_1+c_2(1-Y_i)+c_3V_i] + \tau^{-1}c_3EW_\tau \\ &= \tau^{-1}\sum_1^\infty E\, X_i[c_1+c_2(1-Y_i)+c_3V_i] + \tau^{-1}c_3EW_\tau . \end{aligned} \tag{4.9}$$

Now

$$\begin{aligned} E[X_i(1-Y_i)] &= E\{E[X_i(1-Y_i)|Y_1,\ldots,Y_{i-1},T_1,\ldots,T_i]\} \\ &= E\{X_iE[(1-Y_i)|Y_1,\ldots,Y_{i-1},T_1,\ldots,T_i]\} \\ &= E\{X_i(1-e^{-\lambda T_i})\} , \end{aligned}$$

and similarly,

$$E\,X_iV_i = E\,X_i\{T_i-\lambda^{-1}(1-e^{-\lambda T_i})\}$$

$$E\,W_\tau = E\{\tau - \sum_1^\infty X_iT_i-\lambda^{-1}(1-e^{-\lambda(\tau - \sum_1^\infty X_iT_i)})\} .$$

Substituting into (4.9), we find

$$C(I,\lambda,\tau) = c_3+\tau^{-1}\sum_1^\infty E\,X_i[c_1+d_\lambda(1-e^{-\lambda T_i})] - c_3\lambda^{-1}\tau^{-1}E(1-e^{-\lambda(\tau - \sum_1^\infty X_iT_i)}) . \tag{4.10}$$

If $\lambda \geq c_3/(c_1+c_2)$, then $c_1+d_\lambda \geq 0$, which implies that $c_1+d_\lambda(1-e^{-\lambda T_i}) \geq 0$. Hence, dropping the second term on the right in (4.10), we obtain

$$C(I,\lambda,\tau) \geq c_3-c_3\lambda^{-1}\tau^{-1}$$

and (4.5) follows by letting $\tau \to \infty$. In the remaining case $\lambda < c_3/(c_1+c_2)$ which implies $c_1+d_\lambda < 0$. The function $T^{-1}[c_1+d_\lambda(1-e^{-\lambda T})]$ is minimized by taking $T = T^*_\lambda$, and, using (4.7), this minimum value becomes

$$T^{*-1}_\lambda[c_1+d_\lambda(1-e^{-\lambda T^*_\lambda})] = (c_1+d_\lambda)\lambda(1+\lambda T^*_\lambda)^{-1} < 0 . \tag{4.11}$$

From (4.10)

$$C(I,\lambda,\tau) \geq c_3+T^{*-1}_\lambda[c_1+d_\lambda(1-e^{-\lambda T^*_\lambda})]\tau^{-1}\sum_1^\infty ET_iX_i-c_3\lambda^{-1}\tau^{-1}$$

and using the inequality

$$\tau^{-1}\sum_1^\infty EX_iT_i = E\tau^{-1}\sum_1^\infty X_iT_i = E\tau^{-1}\sum_1^{N_\tau} T_i \leq 1$$

together with (4.11) we have

$$C(I,\lambda,\tau) \geq c_3+T^{*-1}_\lambda[c_1+d_\lambda(1-e^{-\lambda T^*_\lambda})] - c_3\lambda^{-1}\tau^{-1} \tag{4.12}$$

from which (4.5) follows by letting $\tau \to \infty$.

Equality in (4.5) is achieved if $T_i = T^*_\lambda$ a.s. for all i, so that, if λ were known this would constitute an optimal inspection plan. A few properties of T^*_λ as a function of λ are worth noting.

T^*_λ is continuous for $0 < \lambda < c_3/(c_1+c_2)$ and $T^*_\lambda \to +\infty$ as $\lambda \to 0$ or $c_3/(c_1+c_2)$. Also T^*_λ is bounded below by a positive constant; in fact, since

$$e^{-\lambda T^*_\lambda} \geq 1-\lambda T^*_\lambda$$

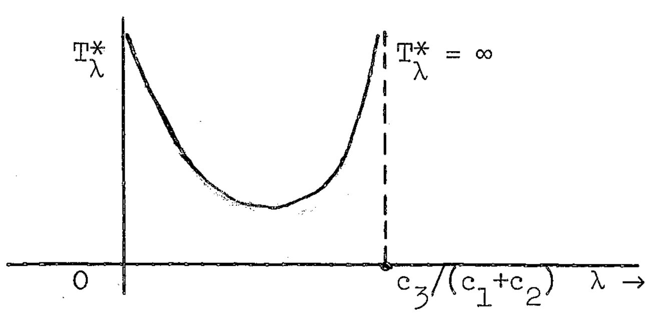
it follows from (4.7) that

$$1-\lambda^2 T^{*2}_\lambda \leq 1 + c_1 d_\lambda^{-1}$$

i.e.,

$$T^*_\lambda \geq (-c_1\lambda^{-2}d_\lambda^{-1})^{1/2} \geq c_1 c_3^{-1} \tag{4.13}$$

for $\lambda < c_3/(c_1+c_2)$. Since, for $\lambda \geq c_3/(c_1+c_2)$, $T^*_\lambda = \infty$ (4.13) holds for all λ. The general form of T^*_λ is indicated in the following figure



Intuitively speaking, when λ is small, the system fails infrequently so that one would expect that the best inspection plan would require infrequent inspections; that is in accordance with the behavior of T^*_λ as $\lambda \to 0$. Also, if λ becomes large, the system would be failing at a high rate. If this rate is high enough, one would expect that it would be more economical to abandon the system and sustain the cost c_3 per unit time rather than to try to maintain the system. Again, this is in accordance with the behavior of T^*_λ as λ increases and, in fact, the critical value is $\lambda = c_3/(c_1+c_2)$.

Now we turn to the case where λ is unknown.

<u>Definition</u>: An inspection plan I is said to be adaptive (relative to $C(I,\lambda)$) if

(4.14) $$C(I,\lambda) = c_3 + T^{*-1}_\lambda [c_1+d_\lambda(1-e^{-\lambda T^*_\lambda})] \ .$$

Next we exhibit and discuss two adaptive plans.

(i) <u>A Maximum Likelihood Plan</u>.

Let $\{s_n\}$ be an increasing sequence of positive numbers such that

$$s_n \to \infty \quad \text{as} \quad n \to \infty$$

and

(4.15) $$\liminf_{n\to\infty} n^{-\alpha} \sum_1^n e^{-\lambda s_i} > 0$$

for all $\lambda \epsilon (\underline{\lambda}, \overline{\lambda})$ and some $\alpha \epsilon (1/2, 1)$. As an example of a sequence satisfying these conditions, we mention

$$s_n = s + \beta \log n$$

where $s > 0$ and $\beta (> 0)$ is such that $\overline{\lambda}\beta + \alpha \leq 1$.

Now we define the following plan. Take $T_1 = s_1$; after n steps compute the maximum likelihood estimate λ_n of λ and then take

(4.16) $$T_{n+1} = \min\{s_{n+1}, T^*_{\lambda_n}\} \ .$$

We shall denote this plan again by ML but it should be noted that it differs somewhat in spirit from the ML plan of section 3. The differences are due to the fact that we have no a priori upper bound on T^*_λ. In fact, if we had defined $T_{n+1} = T^*_{\lambda_n}$ then $\lambda_n \geq c_3/(c_1+c_2)$ for some n (which has positive probability of occurring even if $\lambda < c_3/(c_1+c_2)$) would imply $T_{n+1} = \infty$, resulting in no further inspections and, in general, non-adaptivity of the inspection plan. Hence it is necessary to control the rate at which $\{T_n\}$ can increase and this is the purpose of the sequence $\{s_n\}$.

Theorem 4.2: The ML plan is adaptive.

Proof. First we prove again that $\{\lambda_n\}$ is a strongly consistent sequence of estimates of λ, i.e.,

$$\lambda_n \to \lambda \quad \text{a.s.} \quad \text{as} \quad n \to \infty \ .$$

This result is proved by the following slight modifications of the argument given in the proof of Theorem 3.2. By exactly the same reasoning as that leading to (3.17) we obtain

$$n_k^{-\alpha}\, \xi_{n_k}(\omega) \to 0 \quad \text{as} \quad k \to \infty \ . \tag{4.17}$$

The lower bound on the first two terms on the right of (3.16) given by (3.18) is replaced by

$$\alpha(\mu-\lambda)^{-1} \sum_1^{n_k} e^{-\lambda s_i} \ , \tag{4.18}$$

in which we have used the fact that $c_1 c_3^{-1} \leq T_i \leq s_i$ which follows from (4.13) and (4.16).

Hence, dividing (3.16) by n_k^α and letting $k \to \infty$ while using (4.17) and (4.18) together with (4.15) we again obtain a contradiction. It follows that $\lambda_n \to \lambda$ a.s. as in Theorem 3.2. This implies that

$$T_n \to T^*_\lambda \quad \text{a.s.} \quad \text{as } n \to \infty$$

and this in turn implies that

$$\tau^{-1} N_\tau \to T^{*-1}_\lambda \text{ a.s.}$$

and $\tau^{-1} \sum_1^{N_\tau} (1-e^{-\lambda T_i}) \to T^{*-1}_\lambda (1-e^{-\lambda T^*_\lambda})$ a.s. as $\tau \to \infty$.

Since $\tau \geq \sum_1^{N_\tau} T_i \geq N_\tau c_1 c_3^{-1}$ we have

$$\tau^{-1} N_\tau \leq c_3 c_1^{-1}$$

and

$$\tau^{-1} \sum_1^{N_\tau} (1-e^{-\lambda T_i}) \leq \tau^{-1} N_\tau \leq c_3 c_1^{-1}.$$

Hence by the dominated convergence theorem

$$E\tau^{-1} N_\tau \to T^{*-1}_\lambda$$

and

$$E\tau^{-1} \sum_1^{N_\tau} (1-e^{-\lambda T_i}) \to T^{*-1}_\lambda (1-e^{-\lambda T^*_\lambda})$$

as $\tau \to \infty$. By (4.10),

$$C(ML, \lambda, \tau) = c_3 + c_1 E\tau^{-1} N_\tau + d_\lambda E\tau^{-1} \sum_1^{N_\tau} (1-e^{-\lambda T_i})$$
$$- c_3 \lambda^{-1} \tau^{-1} E(1-e^{-\lambda(\tau - \sum_1^{N_\tau} T_i)})$$

and letting $\tau \to \infty$ it follows from the results above that (4.14) holds.

(ii) <u>A Stochastic Approximation Plan</u>.

Unfortunately T^*_λ cannot be characterized independently of the unknown λ as was possible for T_λ. Consequently, construction of suitable inspection plans using stochastic approximation methods becomes more difficult. We shall discuss such a plan using a sequence of estimates of λ based on a stochastic approximation method closely related to that used to estimate T_λ in section 3. The relation between these methods will be indicated below.

Let $\{s_n\}$ again be an increasing sequence of numbers such that, for all $\lambda \epsilon (\underline{\lambda}, \overline{\lambda})$

(4.19)
$$\begin{cases} \sum_n n^{-2} e^{2\lambda s_n} < \infty \\ \text{and} \\ s_n \to \infty \quad \text{as} \quad n \to \infty . \end{cases}$$

An example of a sequence satisfying these conditions is again

$$s_n = s + \beta \log n$$

where $s > 0$ and β is such that $0 < 2\beta \overline{\lambda} < 1$.

Now, let $T_1 = s_1$ and let λ_1 be arbitrary in the interval $(\underline{\lambda}, \overline{\lambda})$. Having defined $T_1, \ldots, T_n$, $\lambda_1, \ldots, \lambda_n$ we let

$$\lambda_{n+1} = \max\{\underline{\lambda}, \min\{\overline{\lambda}, \lambda_n - n^{-1} B_n (Y_n - e^{-\lambda_n T_n})\}\}$$

and

(4.20)
$$T_{n+1} = \min\{s_{n+1}, T^*_{\lambda_{n+1}}\}$$

where

$$B_n = T_n^{-1} e^{\lambda_n T_n} .$$

We shall denote this plan by SA. In order to see the relation between this plan and the SA plan of section 3 defined by (3.21), let us ignore, for the moment, the truncation of $\{T_n\}$ to the interval $[T_{\overline{\lambda}}, T_{\underline{\lambda}}]$.

Then (3.21) gives approximately

$$T_{n+1} = T_n + n^{-1}A_n(Y_n - e^{-\lambda T_\lambda}) .$$

Taking inverses, we get, approximately

$$T_{n+1}^{-1} = T_n^{-1} + n^{-1}A_nT_n^{-2}(Y_n - e^{-\lambda T_\lambda}) .$$

Since $\lambda_n = -T_n^{-1} \log p$, we then have approximately

$$\lambda_{n+1} = \lambda_n + n^{-1}A_nT_n^{-2}(\log p)(Y_n - e^{-\lambda T_\lambda}) . \tag{4.21}$$

Since $\log p = -\lambda T_\lambda$ and $A_n = p^{-1}\lambda_n^{-1} = \lambda_n^{-1}e^{\lambda T_n}$,

$$A_nT_n^{-2}\log p = -(\lambda\lambda_n^{-1})(T_\lambda T_n^{-1})T_n^{-1}e^{\lambda T_\lambda} . \tag{4.22}$$

Comparing (4.21) and (4.22) we see that the two iterative relations defining $\{\lambda_n\}$ differ only in that λ and T_λ in (4.21) have been replaced by the estimates λ_n and T_n.

Theorem 4.3: The plan SA is adaptive.

Proof. It suffices again to show that $\lambda_n \to \lambda$ a.s. as $n \to \infty$. First we let

$$W_n = n^{-1}B_n(Y_n - e^{-\lambda T_n}).$$

Then

$$E(W_n | W_1, \ldots, W_{n-1}) = 0$$

and

$$\begin{aligned} \text{Var}(W_n) &= n^{-2}E\{B_n^2 e^{-\lambda T_n}(1-e^{-\lambda T_n})\} \\ &\leq n^{-2}E\{T_n^{-2}e^{2\lambda_n T_n}e^{-\lambda T_n}\} \\ &\leq n^{-2}c_3^2c_1^{-2}e^{2\lambda s_n} . \end{aligned}$$

By (4.19)

$$\sum_n \mathrm{Var}(W_n) < \infty ,$$

and it follows again from Theorem D, p. 387 of [7] that $\sum W_n$ converges a.s.. Further,

$$n^{-1}B_n \leq n^{-1}c_3c_1^{-1}e^{\lambda s_n} \to 0 \quad \text{as} \quad n \to \infty$$

and, for $\lambda_n - \lambda \geq \epsilon$,

$$\begin{aligned} &\sum n^{-1}B_n(e^{-\lambda T_n} - e^{-\lambda_n T_n}) \\ &= \sum n^{-1}B_n e^{-\lambda_n T_n}(e^{(\lambda_n - \lambda)T_n} - 1) \\ &\geq \sum n^{-1}B_n e^{-\lambda_n T_n} \epsilon\, T_n \\ &\geq \epsilon \sum n^{-1} \\ &= \infty . \end{aligned}$$

With these three facts in hand an analog of the proof of theorem 3.3 can be given to show that $\lambda_n \to \lambda$ a.s. as $n \to \infty$. We shall not give the details again.

One possible way to compare adaptive plans is, as in section 3, to compare the asymptotic distributions of T_n and choose that with the smallest asymptotic variance.

Theorem 4.4: For both the ML and the SA plans we have

(i) if $\lambda < c_3/(c_1+c_2)$ then

$$\sqrt{n}(T_n - T_\lambda^*) \xrightarrow{\mathcal{L}} N(0, (e^{\lambda T_\lambda^*} - 1)(T_\lambda^{*-1} \frac{dT_\lambda^*}{d\lambda})^2)$$

and

$$\sqrt{n}(\lambda_n - \lambda) \xrightarrow{\mathcal{L}} N(0, (e^{\lambda T_\lambda^*} - 1)T_\lambda^{*-2}) ;$$

(ii) <u>if</u> $\lambda > c_3/(c_1+c_2)$ <u>then, with probability 1, for all</u> n <u>sufficiently large,</u>

$$T_n = s_n \ .$$

<u>Proof</u>. (i) For the first part the proof is quite analogous to that of Theorem 3.4 and we omit the details.

(ii) Since $\lambda_n \to \lambda > c_3/(c_1+c_2)$ a.s., we have, for all n sufficiently large,

$$T^*_{\lambda_n} = \infty \quad \text{i.e.,} \ T_n = \min(s_n, T^*_{\lambda_n}) = s_n \ .$$

This theorem indicates that these two plans are again asympototically equivalent and the SA plan seems to be preferable only in that it involves somewhat simpler computations. It is still true that the SA plan requires the calculation of T^*_λ, thus requiring the solution of equation (4.7) at each step. It is possible to introduce a plan that will also simplify this calculation by using an approximation method.

At present it is not known whether the two plans considered have an optimum property of the type possessed by the equivalent plans in the previous section as indicated by Theorem 3.5.

5. <u>Extensions</u>.

We are now considering the extension of these results along the following lines:

(i) Replacement of the exponential distribution by a more general failure distribution.

(ii) Other types of cost functions.

(iii) Constraints on the inspection times.

This list obviously does not exhaust all possibilities.

REFERENCES

[1] Albert, A.E. (1961), "The sequential design of experiments for infinitely many states of nature," Ann. Math. Statist. 32, 774-799.

[2] Barlow, R. and Hunter, L. (1960), "Optimum checking procedures," Proc. Seventh Nat'l. Symp. on Rel. and Qual. Cont., No. 9, 73.

[3] Chernoff, H. (1959), "Sequential design of experiments," Ann. Math. Statist. 30, 755-770.

[4] Hunter, L. (1963), Optimum checking procedures, Statistical Theory of Reliability, Edited by M. Zelen, University of Wisconson Press.

[5] Kullback, S, (1959), Information theory and statistics, John Wiley, New York.

[6] LeCam, L. (1953), "On some asymptotic properties of maximum likelihood estimates and related Bayes estimates," Univ. of Calif. Publs. Statistics, 1, 277-329.

[7] Loeve, W. (1960), Probability Theory, Second edition.

[8] Robbins, H. and Monro, S. (1951), "A stochastic approximation method," Ann. Math. Statist. 22, 400-407.

[9] Sacks, J. (1958), "Asymptotic distribution of stochastic approximation procedures," Ann. Math. Statist. 29, 373-405.

[10] Venter, J.H. (1963), "On stochastic approximation methods," Doctoral dissertation, Univ. of Chicago.

[11] Wald, A. (1949), "Note on the consistency of the maximum likelihood estimates," Ann. Math. Statist. 20, 595-601.