ON SOME SEARCH STRATEGIES IN RELIABILITY ESTIMATION UNDER RANDOM CENSORSHIP

Shanti S. Gupta
Purdue University

and

Tommaso Gastaldi
Università degli Studi
di Roma "La Sapienza"
and Purdue University

Technical Report #94-12C

Department of Statistics
Purdue University

May 1994

△

ON SOME SEARCH STRATEGIES IN RELIABILITY ESTIMATION UNDER RANDOM CENSORSHIP*

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907-1399

Tommaso Gastaldi
Dipartimento di Statistica, Probabilità e Statistiche Applicate
Università degli Studi di Roma "La Sapienza", 00185 - Roma, ITALIA
and
Department of Statistics
Purdue University
West Lafayette, IN 47907-1399

Abstract

A statistical study on inspection strategies in life-testing for multi-component systems with incomplete information on the cause of failure is carried out. In particular, optimal or nearly optimal inspection strategies are discussed, which allow the user to obtain substantial savings. Simulation results are presented on the respective performance of these strategies.

Key words: random censoring, masked data, multi-component system, competing risk model.

1. INTRODUCTION

This paper is primarily concerned with some techniques to improve the efficiency of standard estimation methods within some frameworks of interest to reliability and biometry scientists. We will assume the reliability point of view to illustrate the practical aspects

*This research was supported in part by NSF Grant DMS-8923071 at Purdue University.

of our work. In particular, we demonstrate the importance of a careful planning of the component inspection in life-tests for multi-component systems, where an incomplete (censored) search for the cause of failure is carried out. Also, simulation results to help the user devise an efficient inspection strategy are provided. In fact, an appropriate selection of the inspection strategy may significantly shorten the time devoted to the inspection itself and yield important savings in the component inspection process. This may also mean that, with a given budget, we might increase the size of our experiment and thus, hopefully, attain a better precision in the estimation.

Before proceeding, we need to introduce the model and some notations. The framework we consider may be viewed as a generalized version of the well-known random censorship model, the generalization consisting of the following. First, we assume several independent concurrent censoring variables instead of only one. These variables are subject to mutual censoring, in the sense that any of the variables can be censored by any one of the other. Second, we assume that the information about the variable which actually generated the observed data, and hence censored the observation of the remaining variables, is possibly incomplete, in a sense that will soon be made precise. We shall now explain in detail the above concepts and, then, illustrate the reason of this extension.

This paper is organized as follows. In Section 2, we present our model, at the same time showing an application to reliability. Section 3 contains the formulation of the best inspection strategy problem from a decision-theoretic point of view and deals with some best (or nearly-best) inspection strategies. Section 4 is concerned with an experimental evaluation of the performance of the proposed strategies. Section 5 closes the article with some discussion and concluding remarks.

2. MODEL

Let us begin with the classical random censorship model. Denote by $\{T_{i1}\} \equiv (T_{11}, \dots, T_{n1})$, and $\{T_{i2}\} \equiv (T_{12}, \dots, T_{n2})$ two sequences of n independent and identically distributed (i.i.d.) random variables (r.v.'s) with the corresponding distributions $F_1(t)$, and $F_2(t)$. Let the two sequences $\{T_{i1}\}$ and $\{T_{i2}\}$ be independent of each other. Set $\delta_i \equiv I(T_{i1} \leq T_{i2})$ and $T_i \equiv \min\{T_{i1}, T_{i2}\}, i = 1, \dots, n$, where $I(A)$ denotes the indicator function of the event A. It is assumed, in the random censorship model, that the observed data consists of the *censored sample* (T_i, δ_i), and that the original sequences $\{T_{i1}\}$ and $\{T_{i2}\}$ are not observed. If further independent censoring variables, $T_{i3}, \dots, T_{ir}$, were now introduced, we could still use the above model whenever our interest lies in making inferences solely on the variable T_{i1}. In fact, the role early played by T_{i2} would be simply replaced by $Z_i \equiv \min\{T_{i2}, \dots, T_{ir}\}$. The point of view we consider consists, instead, in assuming a symmetrical role for the r variables, in the sense that we are interested in making *simultaneous* inferences on their reliabilities. This is made possible by the assumption that the data also contains information on the particular variable which censores the remaining variables. This information will also be assumed to be "fuzzy," in the sense that, for each i, the exact knowledge of the censoring variable is not guaranteed, while what is assumed to be known is a set of possible censoring variables, among which there certainly is the variable that was actually observed. We shall now illustrate the foregoing ideas by an example. We will sometimes refer to the terminology introduced in the following example to illustrate a practical interpretation of our theory. However, our development has a general appeal and can be applied as well to other type of problems, like, for instance, biological survival studies. A formal description of the model follows the example.

Example. Consider a number n of r-out-of-r systems $\Sigma_1, \dots, \Sigma_n$ which are being tested, each system consisting of $r, r \geq 3$, components (or modules) $\mathcal{C}_1, \dots, \mathcal{C}_r$ in series. The random lifelength of component $\mathcal{C}_j, j \in \{1, \dots, r\}$, in system $\Sigma_i, i \in \{1, \dots, n\}$, is denoted by T_{ij}. The rn random variables (r.v.'s) $T_{ij}, i = 1, \dots, n, j = 1, \dots, r$, are assumed to be *independent*. For *each* $j \in \{1, \dots, r\}$, the r.v.'s $T_{1j}, \dots, T_{nj}$, each of which represents the lifelength of the component $\mathcal{C}_j$ in system Σ_i, are also assumed to be *identically distributed*, with a common distribution $F_j(t; \theta_j)$. The random lifelength of each system Σ_i is defined as the first order statistic from the sample $\{T_{i1}, \dots, T_{ir}\}$ and denoted by T_i. For each $i \in \{1, \dots, n\}$, after failure of the system Σ_i at time t_i, a possibly censored analysis on the causes of failure may be carried out and, as a result, a non-empty set, say s_i, is isolated which certainly contains the subscript of the failed component. The incompleteness of the analysis on the cause of failure may be due to several factors. For instance, to minimize downtime, an entire multi-component card or module could be replaced in a computer system without doing further analysis to determine exactly the failed component. Other reasons usually include economical considerations, or technology or time constraints.

The above model is a generalization of the so called *competing risk model*, where, less generally, it is assumed that the component which actually failed is exactly known. Early references on competing risks are Mendenhall and Hader (1958) and Cox (1959). For surveys or further references on such a model, see, for instance, David (1974), David and Moeschberger (1978), and Basu and Klein (1982). A recent paper dealing with Weibull distributed competing lifelengths in a Bayesian framework is Berger and Sun (1993).

In particular, the type of incompleteness we are concerned with is sometimes referred to, in the reliability literature, as "masking," in consideration of the fact that the actual cause of failure may be "masked" by other possible causes. Fundamental references are Miyakawa (1984), Usher and Hodgson (1988), Guess, Usher, and Hodgson (1991). For a good review on the problem of masked system life data, see Usher (1993).

A more general model, including as particular cases the frameworks considered in the present article, is that with *partially classified* data (Gastaldi and Gupta, 1994).

Multivariate censorship model with incomplete data. Let $T_{1j}, \ldots, T_{nj}, j = 1, \ldots, r$, be r *independent* sequences of i.i.d. random variables with the corresponding distribution $F_j(t; \theta_j)$ indexed by the (vector) parameter θ_j.

Denote $I \equiv \{1, \ldots, n\}$, $J \equiv \{1, \ldots, r\}$, and for any $j \in J$ and $t > 0$, let $\overline{F}_j(t; \theta_j) \equiv 1 - F_j(t; \theta_j)$ be the *reliability* (or *survival*) function of $F_j(t; \theta)$ at time t. By $\lambda_j(t; \theta_j) \equiv f_j(t; \theta_j)/\overline{F}_j(t; \theta_j)$ denote the *failure rate* (or *hazard*) function. Also, for each $i = 1, \ldots, n$, let j_i^* be the subscript in $\{1, \ldots, r\}$ such that $T_{ij_i^*} = \min\{T_{i1}, \ldots, T_{ir}\}$ and denote $T_i \equiv T_{ij_i^*}$. It is assumed that j_i^* is *unique* with probability 1.

The observed data consists of the following set

$$\mathcal{D} \equiv \{(t_1, s_1), (t_2, s_2), \ldots, (t_n, s_n)\} \tag{2.1}$$

where t_i is a realization of the r.v. T_i and s_i is a subset of J such that $j_i^* \in s_i$, referred to as the *set of possible causes of failure*.

The above model contains a large class of practical frameworks obtainable by specifying the actual process which gives rise to the sets s_i. There are, in fact, several situations which may cause incompleteness of the data. We will focus on the case when the sets s_i are formed through a censored search for the cause of failure.

Truncated search for the censoring variable. Here we continue describing our model formally; a practical interpretation of the new definitions is shown in the following paragraph.

For convenience of notation, we use the following generalized indicator function:

$$\text{If}[E; r_1, r_2] \equiv \begin{cases} r_1 & \text{if event } E \text{ is true} \\ r_2 & \text{if event } E \text{ is false} \end{cases} \tag{2.2}$$

where r_1, r_2, can be objects of any type.

Given a realization t_i of the r.v. T_i, let $\alpha_{ij}(t_i)$ and $\beta_i(t_i)$ be two given nonnegative functions and $\omega_i(t_i) \equiv < j_{i1}, \ldots, j_{ir} >$ a mapping of $[0, \infty)$ onto the set of all *permutations* of the elements of J. (We use angle brackets to mean that the order is taken into consideration.)

Denote by C_i the, possibly empty, random set of all j_{ih} in $\omega_i(t_i), h = 1, \ldots, r$, satisfying

$$(\sum_{s=1}^{h} \alpha_{ij_{is}}(t_i) \leq \beta_i(t_i)) \text{ and } (h \leq r - 1) \tag{2.3}$$

taken in the same order as they appear in $\omega_i(t_i)$. Let $\nu_i(t_i)$ indicate the number of subscripts in $\omega_i(t_i)$ satisfying (2.3). Clearly, $\nu_i(t_i)$, in general, is a function of $\omega_i(t_i), \beta_i(t_i)$, and $\alpha_{i1}(t_i), \ldots, \alpha_{ir}(t_i)$; however, for simplicity, these dependences will be disregarded in our notation.

Consider the following decomposition of $\omega_i(t_i)$:

$$\omega_i(t_i) = C_i + \overline{C}_i \tag{2.4}$$

where $C_i \equiv \text{If } [\ (\nu_i(t_i) = 0); \emptyset, < j_{i1}, \ldots, j_{i\nu_i(t_i)} >\ ], \overline{C} \equiv < j_{i\nu_i(t_i)+1}, \ldots, j_{ir} >$, and "+" is a symbol for concatenation.

The random set S_i is defined as

$$S_i \equiv \text{If}[\ (j_i^* \in C_i); \{j_i^*\}, \overline{C}_i\ ] \tag{2.5}$$

while by s_i we refer to a realization of S_i. Clearly, the cardinality of S_i is either 1 or $r - \nu_i(t_i)$.

The likelihood function of the parameters given the observed data (2.1) is

$$L(\theta_1, \ldots, \theta_r | \mathcal{D}) = \prod_{i \in I} \{ \sum_{j \in s_i} \lambda_j(t_i; \theta_j) \prod_{w \in J} \overline{F}_w(t_i; \theta_w) \}. \tag{2.6}$$

The practical meaning of the above definitions will now be illustrated.

Reliability interpretation. With reference to our previous example, assume that, after the failure of each system $\Sigma_i, i = 1, \ldots, n$, a sequential search for the failed component within the system is carried out by checking one by one the system components, taken in the same order as their subscripts appear in the ordered set $\omega_i(t_i) \equiv < j_{i1}, \ldots, j_{ir} >$. Hereafter, we will also refer to such a set as the *(component) inspection strategy*. Assume that $\alpha_{ij}(t_i)$ represents the time needed to carry out the failure analysis on component C_j in system Σ_i, also referred to as the *component checking time*. Let $\beta_i(t_i)$ denote the maximum length of time available for the failure analysis on system Σ_i, also referred to as the *system checking time limit*. The random set C_i is formed of all and only those subscripts of the components that can be inspected, given the time constraint $\beta_i(t_i)$. By assumption, the cardinality of C_i is *at most* equal to $r-1$, because the inspection of $r-1$ components ensures the determination of the cause of failure. The set s_i is formed of all and only those *possible causes of failure* isolated after a failure analysis conducted, for a period *not longer* than $\beta_i(t_i)$, on the r components of system Σ_i, each of which requires a time *exactly* equal to $\alpha_{ij}(t_i)$ to be inspected. This can be viewed as a special model with masked data. In particular, it can be shown that the estimators proposed by Usher and Hodgson (1988) are the MLEs in this particular context where the masking arises from a censored search for the cause of failure, while in general they are not.

3. BEST STRATEGY SELECTION PROBLEMS AND OPTIMAL STRATEGIES

The main concern one has when estimating reliabilities within a model with incomplete search for the cause of failure is that of selecting the best component inspection strategy in order to minimize the global amount of masking and, hence, attain the highest precision in the estimation process. It is, in fact, intuitive that the order in which the components are processed in the failure analysis affects the chance of isolating the actual cause of failure of a system Σ_i within the given checking time limit $\beta_i(t_i)$. Furthermore, the greater the indetermination about the cause of failure, the lesser the precision of any consistent estimate of the component reliabilities. It is, therefore, natural, to look for a component inspection strategy $\omega_i(t_i)$ which minimizes the chance of extensive masking. The idea of increasing the precision of the failure rates estimates through a meaningful inspection plan in the context of masked data was sketched in Gastaldi (1993). Here, it is developed and substantiated by an extensive simulation study.

We consider the following decision-theoretic setup. Define the set of all possible decision Δ_r as the set of all permutations of the elements of J. Given a failure time t_i, and an inspection strategy $\omega_i(t_i) \equiv < j_{i1}, \ldots, j_{ir} >$ in Δ_r, the system checking time limit $\beta_i(t_i)$ will suffice to inspect at most a number of components equal to $\nu_i(t_i)$. With each strategy is associated a certain loss which also depends on the other parameters in the model. We consider the following loss functions.

Loss functions. The first loss function we consider is the following:

$$L_1(S_i, \theta_1, \ldots, \theta_r, \alpha_{i1}, \ldots, \alpha_{ir}, \beta_i, \omega_i) \equiv \text{If}[\ (|S_i| = 1); 0, 1\ ] \qquad (3.1)$$

i.e., a zero loss when the cause of failure is isolated and a loss equal to 1 (or any other constant) when the cause of failure is not found, because the inspection strategy was such that the components checked in the available time $\beta_i(t_i)$ were all different from that which caused the system failure.

The above loss function may sometimes be rather unrealistic in that it does not take into account economical factors such as, for instance, the cost of the inspection process per unit of time. By assuming inspection costs proportional to the component checking times, a loss function which would incorporate such issues is the following:

$$L_2(S_i, \theta_1, \ldots, \theta_r, \alpha_{i1}, \ldots, \alpha_{ir}, \beta_i, \omega_i) \equiv \sum_{h=1}^{\nu_i(t_i)} \alpha_{ij_{ih}}(t_i) I(h \leq w^*) \qquad (3.2)$$

with $\nu_i(t_i) \geq 1$, where w^* denotes the position of the failed component in the inspection sequence $< j_{i1}, \ldots, j_{ir} >$, i.e., a loss proportional to the time employed in the search process.

It may be worth noticing that loss function (3.1) is not a special case of loss (3.2), and that it cannot be obtained, as one might suspect, by placing constant checking times in

(3.2), since the latter, even with constant checking times, would also take into account the number of components inspected to find the failed component, while (3.1) does not. Furthermore, we might already expect that since loss (3.1), does not take into consideration the length of the inspection process, yields several best strategies. In fact, given a best strategy $< j_{i1}, \ldots, j_{ir} >$, every other strategy $< j_{i1}^{(1)}, \ldots, j_{i\nu_i(t_i)}^{(1)} > + < j_{i\nu_i(t_i)+1}^{(2)}, \ldots, j_{ir}^{(2)} >$, where $< j_{i1}^{(1)}, \ldots, j_{i\nu_i(t_i)}^{(1)} >, < j_{i\nu_i(t_i)+1}^{(2)}, \ldots, j_{ir}^{(2)} >$ denote any two permutations of $< j_{i1}, \ldots, j_{i\nu_i(t_i)} >$ and $< j_{i\nu_i(t_i)+1}, \ldots, j_{ir} >$, respectively, will also be optimal.

Selection problem. Our selection goal is defined as follows. Determine the inspection strategies, in the set Δ_r, such that the following conditional average loss (or *risk*) is minimum:

$$E[L(S_i, \theta_1, \ldots, \theta_r, \alpha_{i1}, \ldots, \alpha_{ir}, \beta_i, \omega_i) \mid T_i = t_i]. \tag{3.3}$$

Such strategies will be referred to as best. We shall now give some results on the best inspection strategies, under the most usual practical frameworks.

Optimal strategies. We can restrict our attention to the systems such that there exist at least one strategy $< j_{i1}, \ldots, j_{ir} >$, in Δ_r satisfying $\alpha_{ij_1}(t_i) \leq \beta_i(t_i)$, i.e., there is time to check at least one component:

$$\min_{j=1,\ldots,r} \alpha_{ij}(t_i) \leq \beta_i(t_i) \tag{3.4}$$

which ensures that, at least for one strategy, $\nu_i(t_i) \geq 1$.

Case 1 - Best inspection strategy under loss function (3.1). Notice that, under loss function (3.1), the risk function (3.3) can be interpreted as the probability that the cause of failure is *not* found during the inspection.

In this case, we have that minimizing (3.3) with L replaced by loss (3.1) is equivalent to maximizing

$$E[I(|S_i| = 1) \mid T_i = t_i] \tag{3.5}$$

$$= Pr[\text{One of the inspected components caused the system failure}] \tag{3.6}$$

$$= Pr[\bigcup_{j \in C_i} (j_i^* = j) \mid T = t_i] = \frac{1}{\Lambda(t_i, \theta_1 \ldots, \theta_r)} \sum_{j \in C_i} \lambda_j(t_i, \theta_j). \tag{3.7}$$

where we have denoted by $\Lambda(t_i, \theta_1, \ldots, \theta_r)$ the system failure rate at time t_i, i.e., $\sum_{j \in J} \lambda_j(t_i, \theta_j)$.

Finding a general closed form solution to the above maximization problem, despite its simple formulation, is not always straightforward unless special assumptions are made on the component checking times or on the failure rates. We will, hence, discuss separately the possible situations.

Subcase 1.1 - The time needed to inspect a component is the same for any component and the failure rates are not all equal. This is a very common case in practice. It occurs, for instance, when the components subject to possible breakdown have a similar or equal degree of accessibility in the systems. In such a case, we have that $\nu_i(t_i)$ does not depend on the inspection strategy and it is equal to $\min\{[\beta_i(t_i)/\alpha_i(t_i)], r-1\}$, where $[x]$ denotes the greatest integer less than or equal to x, and $\alpha_i(t_i) \equiv \alpha_{ij}(t_i), j = 1, \ldots, r$. It is easy to see that the optimal strategies $\omega_i(t_i) \equiv < j_{i1}, \ldots, j_{ir} >$ are all and only those such that

$$\textit{for any } \; j_{ic} \in C_i \; \text{ and } \; j_{i\bar{c}} \in \overline{C}_i, \; \text{ we have } \; \lambda_{j_{ic}}(t_i, \theta_{j_{ic}}) \geq \lambda_{j_{i\bar{c}}}(t_i, \theta_{j_{i\bar{c}}}). \tag{3.8}$$

The intuitive meaning of this solution is that, in order to reduce the probability of not determining the failed component, we should inspect the first $\nu_i(t_i)$ components with highest failure rates at the system failure time. Notice that this also minimizes the conditional expected time for component inspection given $T_i = t_i$ if the subscripts in C_i are arranged in nondecreasing order w.r.t. the corresponding $\lambda_j(t_i, \theta_j)$'s.

Subcase 1.2 - The times needed to inspect the components are not all equal and the failure rates are all equal. This is the case when equal or very similar components have different accessibility in a system. In this case, it is straightforward to notice that the optimal strategies $\omega_i(t_i) \equiv < j_{i1}, \ldots, j_{ir} >$ are all and only those such that

$$\textit{for any } j_{ic} \in C_i \; \text{ and } \; j_{i\bar{c}} \in \overline{C}_i, \; \text{ we have } \; \alpha_{ij_{ic}}(t_i) \leq \alpha_{ij_{i\bar{c}}}(t_i). \tag{3.9}$$

In this case, the optimal strategies are quite intuitive, and consist of inspecting first the $\nu_i(t_i)$ components with smallest checking times.

Subcase 1.3 - The times needed to inspect the components are not all equal and the failure rates are not all equal. In this case the nature of the problem does not allow to express the best inspection strategies in a closed form. Intuition suggests that, in this case, the optimal strategies must be a "compromise" between those encountered in the previous two subcases. It should be intuitive that in order to maximize the sum of the failure rates of the components inspected during the available time $\beta_i(t_i)$, we should allocate in C_i the components with higher failure ratio $\lambda_{ij}(t_i, \theta_j)/\alpha_{ij}$ at time t_i, in order to use as efficiently as possible the time spent in the inspection process. Clearly, this is achieved, for instance, by giving higher precedence, in the inspection, to those components with higher ratios of failure rate to inspection time. In the cases of two components with equal ratios, clearly, the precedence should be given to the component with higher failure rate (or, equivalently, lower checking time). Bearing these considerations in mind, it is clear that the following algorithm yields a good approximation to a best inspection strategy.

Algorithm 3.1

Step 0: Set the maximum time available for inspection equal to $\beta_i(t_i)$.

Step 1: Exclude from consideration of inspection all the components whose inspection times exceed the maximum time available for inspection *or* that have already been inspected.

Step 2: Quit the inspection process if the set of the components which have not been excluded for consideration to be inspected is empty.

Step 3: Among the components not excluded from inspection, select and check the one with highest ratio $\lambda_{ij}/\alpha_{ij}(t_i)$. If there are two or more of such components, inspect the one (or one of those) with largest λ_{ij} (or, equivalently, smallest $\alpha_{ij}(t_i)$).

Step 4: Reduce the maximum available time for inspection by the time necessary to inspect the component selected in the previous step.

Step 5: Go to step 1.

Unfortunately, due to the the discrete nature of the component allocation in the sets C_i and $\overline{C}_i$, there is no guarantee that the above are the optimal strategies, although we can expect them to be good approximations. This is due to the fact that there might be some residual time $\beta_i(t_i) - \sum_{j \in C_i} \alpha_{ij}(t_i)$, left at the end of the inspection process, which makes it possible that by replacing some of the checked components by a larger number of uninspected components with smaller ratios $\lambda_{ij}/\alpha_{ij}(t_i)$ and larger sum of failure rates, we might obtain smaller value of the average loss, thanks to the use of part or all of the unutilized time. Strategies generated by Algorithm 3.1 have, however, at least two features that make them suitable for practical purposes: they are generally close to the optima and can be found with a negligible computing effort.

Case 2 - Best inspection strategy under loss function (3.2). In this case, the risk function can be interpreted as the mean time needed to identify the failed component. We have:

$$E[L_2(S_i, \theta_1, \ldots, \theta_r, \alpha_{i1}, \ldots, \alpha_{ir}, \beta_i, \omega_i) \mid T_i = t_i]$$

$$= E[\sum_{h=1}^{\nu_i(t_i)} \alpha_{ij_h}(t_i)\ I(h \leq w^*) \mid T_i = t_i] \tag{3.10}$$

$$= \frac{1}{\Lambda(t_i, \theta_1, \ldots, \theta_r)} \sum_{s=1}^{\nu_i(t_i)} \text{If}[\ (s < \nu_i(t_i)); \lambda_{j_{is}}(t_i, \theta_{j_{is}}), \sum_{v=\nu_i(t_i)}^{r} \lambda_{j_{iv}}(t_i, \theta_{j_{iv}})\ ] \sum_{w=1}^{s} \alpha_{ij_{iw}}(t_i) \tag{3.11}$$

$$= \frac{1}{\Lambda(t_i, \theta_1, \ldots, \theta_r)} \sum_{s=i}^{\nu_i(t_i)} \alpha_{ij_{is}}(t_i) \sum_{w=s}^{r} \lambda_{j_{iw}}(t_i, \theta_{j_{iw}}). \tag{3.12}$$

Let us consider the following three main subcases.

Subcase 2.1 - The component checking times are all equal and the failure rates are not all equal. Denote $\alpha_i(t_i) \equiv \alpha_{ij}(t_i)$ for all $j \in J$. In this case, we have the following risk:

$$\frac{\alpha_i(t_i)}{\Lambda(t_i, \theta_1, \ldots, \theta_r)}[\sum_{s=1}^{\nu_i(t_i)-1} s\lambda_{j_{is}}(t_i, \theta_{j_{is}}) + \nu_i(t_i) \sum_{v=\nu_i(t_i)}^{r} \lambda_{j_{iv}}(t_i, \theta_{j_{iv}})] \tag{3.13}$$

which is minimized by those strategies $\omega_i(t_i) \equiv < j_{i1}, \ldots, j_{ir} >$ such that $\lambda_{j_{ih}}(t_i, \theta_{j_{ih}}) \geq \lambda_{ij_{i,h+1}}(t_i, \theta_{j_{i,h+1}})$, for all $h = 1, \ldots, \nu_i(t_i) - 1$.

Subcase 2.2 - The times needed to inspect the components are not all equal and the failure rates are all equal. Denote $\lambda_j(t_i, \theta_j) \equiv \lambda(t_i)$ for all $j \in J$. In this case, the risk is proportional to the weighted sum of the component checking times of the inspected components, with weights forming a decreasing sequence.

$$\frac{1}{r}\sum_{s=1}^{\nu_i(t_i)} \alpha_{ij_{is}}(t_i)(r-s+1) \tag{3.14}$$

and is minimized by inspecting the system components in such an order $\omega_i(t_i) \equiv\ < j_{i1}, \ldots, j_{ir} >$ as to have $\alpha_{ij_{ih}}(t_i) \leq \alpha_{ij_{i,h+1}}(t_i)$, for all $h = 1, \ldots, \nu_i(t_i) - 1$.

Subcase 2.3 - The times needed to inspect the components are not all equal and the failure rates are not all equal. In this general case, it is not possible to give the best strategies in a closed form. However, while for r small it is always possible to scan the decision space in search of a best strategy, for r large we provide locally best strategies, defined as follows.

Definition. A strategy is said to be *nearly-best* if it is not dominated by any other strategy obtainable from it by swapping any two adjacent elements.

The following theorem helps us to find the above kind of strategies.

Theorem. Any strategy $\omega_i(t_i) \equiv< j_{i1}, \ldots, j_{ir} >$ such that for every two adjacent elements $j_{ih}, j_{i,h+1} \in C_i, h = 1, \ldots, \nu_i(t_i) - 1$, we have

$$\lambda_{ij_{ih}}(t_i, \theta_{j_{ih}})/\alpha_{ij_{ih}}(t_i) \geq \lambda_{ij_{i,h+1}}(t_i, \theta_{j_{i,h+1}})/\alpha_{ij_{i,h+1}}(t_i) \tag{3.15}$$

is *nearly-best*.

Proof. Denote by $\omega_i^*(t_i) \equiv< j_{i1}^*, \ldots, j_{ir}^* >\equiv C_i^* + \overline{C_i^*}$ a strategy such that $\lambda_{ij_{ih}^*}/\alpha_{ij_{ih}^*} \geq \lambda_{ij_{i,h+1}^*}/\alpha_{ij_{i,h+1}^*}$ for every pair $j_{ih}^*, j_{i,h+1}^* \in C_i^*$. Assume that there exists a strategy $\omega_i'(t_i) \equiv\ < j_{i1}', \ldots, j_{ir}' >\equiv\ < j_{i1}^*, \ldots, j_{ih+1}^*, j_{ih}^*, \ldots j_{ij_{\nu_i(t_i)}}^*, \ldots, j_{ir}^* >$ (notice the swapping of the two internal elements j_{ih}^*, j_{ih+1}^*) which yields a lower average loss. Then, by (3.12) we have

$$\frac{1}{\Lambda}\sum_{s=1}^{\nu_i(t_i)} \alpha_{ij_{is}'}(t_i)\sum_{w=s}^{r} \lambda_{j_{iw}'}(t_i, \theta_{j_{iw}'}) < \frac{1}{\Lambda}\sum_{s=1}^{\nu_i(t_i)} \alpha_{ij_{is}^*}(t_i)\sum_{w=s}^{r} \lambda_{j_{iw}^*}(t_i, \theta_{j_{iw}^*}) \tag{3.16}$$

$$\Leftrightarrow \lambda_{ij_{ih}^*}(t_i, \theta_{j_{ih}^*})/\alpha_{ij_{ih}^*}(t_i) < \lambda_{ij_{i,h+1}^*}(t_i, \theta_{j_{i,h+1}^*})/\alpha_{ij_{i,h+1}^*}(t_i) \tag{3.17}$$

that is, a contradiction. □

Notice that Algorithm (3.1) provides a strategy which satisfy (3.15) and is hence *nearly-best*.

A closed-loop problem? Thus far we have treated the failure rates as if they were known. However, the goal itself of the reduction of masking is a shortening of the *estimation process of the failure rates* themselves. This apparent closed-loop problem, finds, in practice, several possible solutions. In fact some preliminary information on the failure rates is often available, for instance from databases of system life data containing information from past experiments conducted on similar systems. If not available, it can be an expression of the *a priori* beliefs of the researcher, supported by the technical knowledge about the components. Finally, in absence of any type of information, it can be obtained, with increasing accuracy, through estimation carried out using the data just observed, in an adaptive process. This is shown in our simulation study. Furthermore, it can be noticed that, in the case of equal checking times, the required information considers only the *order* of the failure rates, and not their magnitudes.

Is there any loss of precision in the attempt to shorten the inspection process? By using an optimizing strategy for the component inspection, we certainly expect significant reduction of masking and time savings in the case of equal component checking times. However, the main concern is whether these savings result in poorer estimates of the reliabilities. In fact, if this were so, we might need more observations to achieve a prespecified precision, and the cost of further experimentation might exceed the savings. An indication on this comes from the work of Guess, Usher and Hodgson (1991), who point out through simulation that, in the three-component case, by reducing the masking, the mean square errors of the MLEs are decreased, in the model with exponentially distributed lifelengths. The simulation in Section 4, conducted on five-component systems, should give the reader an idea of the behavior of the ML estimates under different inspection strategies and thus help choose the best inspection plan in real experiments.

4. SIMULATION STUDY

Here we compare the performances of the Miyakawa (1984) and Usher-Hodgson (1988) MLEs estimators for masked data under different inspection strategies. In particular, we will compare the behaviors of the following kind of strategies:

(1) Fixed inspection order

(2) Reverse of (1)

(3) Random inspection

(4) Increasing checking time

(5) Nearly-best strategy.

The first two strategies simply consist of inspecting the components according to the following orderings $< 1, 2, \ldots, r >, < r, r-1, \ldots, 1 >$, respectively. There is no special reason for picking up these two specific elements from the decision space Δ_r; however they can just be considered as two permutations randomly chosen from Δ_r, since the labels $1, 2, \ldots, r$ have also been randomly assigned to the components. The *random* inspection strategy consists of choosing an inspection order at random from Δ_r, *for each system* $\Sigma_i, i = 1, \ldots, n$. The *increasing checking time* strategy is defined as the sequence $< j_{i1}, \ldots, j_{ir} >$ such that $\alpha_{ij_{ih}}(t_i) \leq \alpha_{ij_{i,h+1}}(t_i), h = 1, \ldots, r-1$. Finally, by *nearly best*, we refer to the strategy generated by Algorithm 3.1.

In order to evaluate the performances of the above inspection strategies, we generated a number m of experiments (life-tests) $\mathcal{E}_h, h = 1, \ldots, m$, of size n and applied these strategies to estimate the failure rates of r components with exponentially distributed lifelengths. We have not considered sequential estimation, since failure rates estimators based on order statistics can be, in general, unavailable. However, when they are available, all results in the paper still apply, provided that the estimators for a fixed sample size are replaced by the estimators based on first order statistics. The reason for this is the fact that the probability $Pr\{j_i = j_i^* \mid T_i = t_i\}$ *does not depend on the position of* Σ_i *within the sequence of systems ordered by their respective failure times.*

For a lack of theoretical development, we have not considered distributions other than the negative exponential. Notice that even in the exponential case the estimation must be generally carried out through numerical methods.

Simulated data

$$\mathcal{D}^{(h)} \equiv \{(t_1^{(h)}, s_1^{(h)}), (t_2^{(h)}, s_2^{(h)}), \ldots, (t_n^{(h)}, s_n^{(h)})\} \tag{4.1}$$

$h = 1, \ldots, m$, for each of the m samples has been obtained as follows. For each system $\Sigma_i^{(h)}, i = 1, \ldots, n$, in the *h-th* sample we have simulated exponential lifelengths for the r components. For instance, in Simulation 1, we have set $n = 500, r = 5, m = 100, \beta_i(t_i) \equiv \beta = 1.05$ for all $i \in I$, and assumed the set of failure rates and component checking times shown in Table 1. It is only in order not to have too many variables that the r component checking times have been taken constant *with respect to failure time and system*, and that the system checking time limits $\beta_i(t_i)$ have been set equal for all systems $\Sigma_i, i = 1, \ldots, n$, and failure times.

Table 1.

Failure rates and component checking times for Simulation 1

	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$
λ_j^{-1} (hours)	85	150	90	190	40
α_{ij} (hours)	0.45	0.25	0.15	0.51	0.5

In Simulation 2, for sake of comparison, we have left everything unchanged but the following parameters: $\beta_i(t_i) \equiv \beta = .75$ for all $i \in I$, and $\alpha_{ij}(t_i) \equiv \alpha = .25$ for all $i \in I, j \in J$, and t_i.

For each life-test $\mathcal{E}_h, h = 1, \ldots, m$, we have generated life and failure of all systems $\Sigma_i^{(h)}, i = 1, \ldots, n$, and finally simulated a censored search of the cause of failure according to the different strategies mentioned above. In each sample $\mathcal{D}^{(h)}, h = 1, \ldots, 1000$, the nearly-best strategy has been defined as coincident with the random strategy up to the 25-th observation. From the 26-th observation on it is computed through Algorithm 3.1, where the λ_j's are replaced by the estimates obtained using the lifelengths already observed. The reason why the use of the nearly-best strategy begins only after some observations is that those are obviously needed to form a preliminary estimate of the failure rates. In fact, we are assuming here that absolutely no information is available on the failure rates λ_j's. If some information were available, it could be used, possibly integrating it with the knowledge coming from the experiment itself, to infer on the ordering of the λ_j's.

Most of the quantities shown in the tables below are defined according to the standard usage. A legend follows; also, a few quantities appearing in the simulation tables, which have not been explicitly defined thus far, are introduced below.

MLEs. The maximum likelihood estimates of the failure rates under the assumption of exponential distributions of the components lifelengths are obtained by numerical solution of the following non-linear ML simultaneous equations:

$$\sum_{\kappa=1}^{2^r-1} \frac{n_{\zeta_\kappa}}{\sum\limits_{w\in\zeta_\kappa}\lambda_w} I(j\in\zeta_\kappa) - \sum_{i=1}^{n} t_i = 0, \quad j\in J, \tag{4.2}$$

where $\zeta_\kappa, \kappa = 1,\dots,2^r-1$, is the sequence of all nonempty subsets of $\{1,\dots,r\}$, taken in any arbitrarily fixed order, and n_{ζ_κ} denotes the number of observations (t_i, s_i) such that $s_i = \zeta_\kappa$. For instance, when $r = 3$ (we show this case for sake of simplicity, even though in our simulation we considered 5 components), we have the following equations:

$$\begin{aligned}
&\frac{n_1}{\lambda_1} + \frac{n_{12}}{\lambda_1+\lambda_2} + \frac{n_{13}}{\lambda_1+\lambda_3} + \frac{n_{123}}{\lambda_1+\lambda_2+\lambda_3} - \sum_{i=1}^{n} t_i = 0 \\
&\frac{n_2}{\lambda_2} + \frac{n_{12}}{\lambda_1+\lambda_2} + \frac{n_{23}}{\lambda_2+\lambda_3} + \frac{n_{123}}{\lambda_1+\lambda_2+\lambda_3} - \sum_{i=1}^{n} t_i = 0 \\
&\frac{n_3}{\lambda_3} + \frac{n_{13}}{\lambda_1+\lambda_3} + \frac{n_{23}}{\lambda_2+\lambda_3} + \frac{n_{123}}{\lambda_1+\lambda_2+\lambda_3} - \sum_{i=1}^{n} t_i = 0.
\end{aligned} \tag{4.3}$$

To solve the system, we have considered two types of iterative algorithms. The first algorithm is given by the following recursive relation:

$$\lambda_j^{(i)} = \phi(n_{\zeta_\kappa}, \kappa = 1,\dots,2^r-1; \lambda_1^{(i-1)},\dots,\lambda_r^{(i-1)}), \quad j = 1,\dots,r, \tag{4.4}$$

$i = 1, 2, \dots, N(\epsilon)$, where ϵ is a positive constant used to define the number $N(\epsilon)$ of iterations. The quantity $N(\epsilon)$ is defined as the index i for which $\cap_{j=1}^{r} |\lambda_j^{(i)} - \lambda_j^{(i-1)}| < \epsilon$, where, in particular, we have set $\epsilon = 10^{-7}$.

The other algorithm is given by the following equations:

$$\begin{aligned}
&\lambda_1^{(i)} = \phi(n_{\zeta_\kappa}, \kappa = 1,\dots,2^r-1; \lambda_1^{(i-1)}, \lambda_2^{(i-1)} \dots, \lambda_r^{(i-1)}) \\
&\lambda_2^{(i)} = \phi(n_{\zeta_\kappa}, \kappa = 1,\dots,2^r-1; \lambda_1^{(i)}, \lambda_2^{(i-1)} \dots, \lambda_r^{(i-1)}) \\
&\dots \\
&\lambda_r^{(i)} = \phi(n_{\zeta_\kappa}, \kappa = 1,\dots,2^r-1; \lambda_1^{(i)}, \dots, \lambda_{r-1}^{(i)}, \lambda_r^{(i-1)})
\end{aligned} \tag{4.5}$$

$i = 1, \dots, N'(\epsilon)$.

This second algorithm differs slightly from the first one in that, as soon as a value for the estimate of $\lambda_j, j = 1, \dots, r-1$, is obtained, it is substituted in the next equations.

As starting values, we have used $\lambda_j^{(0)} = 1$ for all $j = 1, \dots, r$. Since simulation shows that strategy comparison is not particularly affected by the choice of the algorithm used to solve the system, we only show results obtained with the first algorithm. The starting values are also unimportant as far as convergence is concerned; the convergence, generally, takes place after very few iterations, even when using starting values extremely far away from the true values of the failure rates. This was also noticed by Usher and Hodgson (1988) when solving the ML system within the three-component case.

Table of symbols

Term	*Theoretical quantity being estimated*	
Expected values of failure rate estimators	$E(\widehat{\lambda}_j)$	$j = 1, \dots, r$
Average of the above	$(1/r) \sum_{j \in J} E(\widehat{\lambda}_j)$	
Biases of failure rate estimators	$B_j \equiv E(\widehat{\lambda}_j) - \lambda_j$	$j = 1, \dots, r$
Bias ratios	$BR_j \equiv E(\widehat{\lambda}_j)/\lambda_j$	$j = 1, \dots, r$
Geometric average of the above	$(\prod_{j \in J} R_j)^{1/r}$	
Standard deviations	$SD_j \equiv [E(\widehat{\lambda}_j - E(\widehat{\lambda}_j))^2]^{1/2}$	$j = 1, \dots, r$
Relative standard deviations	$RSD_j \equiv SD_j / E(\widehat{\lambda}_j)$	$j = 1, \dots, r$
Average of the above	$(1/r) \sum_{j \in J} RSD_j$	
Square root of Mean square errors	$MSE_j \equiv [E(\widehat{\lambda}_j - \lambda_j)^2]^{1/2}$	$j = 1, \dots, r$
Square root of Relative mean square errors	$RMSE_j \equiv MSE_j / \lambda_j$	$j = 1, \dots, r$
Average of the above	$(1/r) \sum_{j \in J} RMSE_j$	

Defn. 1: Average *Total Masking*. It is the average sum of the size of the sets s_i

observed during the life-test.

$$(1/m)\sum_{h=1}^{m}(\sum_{i=1}^{n} \mid s_i \mid). \tag{4.6}$$

Defn. 2: Average *Mean Masking.* It is the average observed mean of the size of the set S_i.

$$(1/m)\sum_{h=1}^{m} [(1/n)\sum_{i=1}^{n} \mid s_i \mid]. \tag{4.7}$$

Defn. 3: Average *1st Risk Function (Fn.).* It is the average value of the expected loss using loss function (2.1) multiplied by 100. It expresses the average percentage of systems where the cause of failure has not been found.

Defn. 4: Average *Total Inspection Time.* It is the average value of the sum of the inspection times in life-tests.

$$(1/m)\sum_{h=1}^{m} [\sum_{i=1}^{n} IT_i] \tag{4.8}$$

where the inspection time IT_i of system Σ_i is

$$IT_i \equiv \text{ If}[\ (\nu_i(t_i) = 0); 0, \sum_{s=1}^{\nu_i(t_i)} \alpha_{ij_{is}}(t_i)\ I(s \le w^*)\ ] \tag{4.9}$$

(given a strategy $\{j_{i1}, \dots, j_{ir}\}$, we denoted by $j^*_{iw^*}$ the subscript of the failed component and by w^* the position of such subscript within such a strategy) i.e., equal to zero if there is no time to inspect the first component listed in the strategy or else equal to the sum of the checking times of the components inspected within the time $\beta_i(t_i)$ until either the time expires or the failed component is found.

Defn. 5: Average *2nd Risk Fn.* (or *Mean Inspection Time*). This is simply

$$(1/m)\sum_{h=1}^{m} [(1/n)\sum_{i=1}^{n} IT_i]$$

where IT_i is the quantity defined in (4.9).

Defn. 6: Average *"Wasted" Time.* Although the appropriateness of the term can be argued upon, we define as "wasted" time (in the inspection of a system Σ_i) the inspection time minus the time of the failed component if the latter is one of the checked components. In other words, this is the time that is spent in the inspection of components which did not cause the system failure. It will be equal to zero when either the inspection time is zero or $w^* = 1$.

The last column in the tables shows the percentage the wasted time to inspection time.

LIFE TESTS PARAMETERS & STATISTICS

Simulation id. number:	1
Life test size:	500
Starting point for computation of the nearly best strategy:	25
Number of simulated life tests:	100
Average sum of system lifelengths:	8486.733

SYSTEM FEATURES

Number of system components:	5
System checking time limit:	1.05

COMPONENT FEATURES

Component	C_1	C_2	C_3	C_4	C_5
Mean lifelength	85	150	90	190	40
Failure rate	0.01176	0.00667	0.01111	0.00526	0.02500
Inspection time	0.45	0.25	0.15	0.51	0.50

STATISTICS ON THE SIMULATED LIFELENGTHS

Component	C_1	C_2	C_3	C_4	C_5
Average of sample means	84.97	151.05	89.89	190.15	40.11
Average of sample STDs	85.45	151.08	89.62	190.49	39.75

RESULTS ON THE FAILURE RATE ESTIMATORS

MEANS

ESTIMATORS / STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Average (for verification)
Fixed order	0.01180	0.00653	0.01112	0.01478	0.01478	0.01180
Reverse fixed order	0.00982	0.00982	0.00982	0.00495	0.02461	0.01180
Random order	0.01184	0.00648	0.01106	0.00516	0.02446	0.01180
Increasing checking time	0.01180	0.00653	0.01112	0.01478	0.01478	0.01180
Nearly best strategy	0.01028	0.00653	0.01114	0.00674	0.02433	0.01180

BIASES

ESTIMATORS / STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Averages (for verification)
Fixed order	0.0000391	-0.0001342	0.0000131	0.0095136	-0.0102232	-0.0001583
Reverse fixed order	-0.0019446	0.0031535	-0.0012910	-0.0003163	-0.0003933	-0.0001583
Random order	0.0000762	-0.0001827	-0.0000474	-0.0001002	-0.0005376	-0.0001583
Increasing checking time	0.0000391	-0.0001342	0.0000131	0.0095136	-0.0102232	-0.0001583
Nearly best strategy	-0.0014884	-0.0001327	0.0000286	0.0014757	-0.0006748	-0.0001583

OBSERVED BIAS RATIOS E(failure rate' MLE)/(true failure rate)

ESTIMATORS / STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	(Geometric) Averages
Fixed order	1.0033202	0.9798663	1.0011773	2.8075917	0.5910720	1.6333966
Reverse fixed order	0.8347116	1.4730204	0.8838122	0.9399051	0.9842693	1.0053174
Random order	1.0064801	0.9725913	0.9957348	0.9809638	0.9784973	0.9356036
Increasing checking time	1.0033202	0.9798663	1.0011773	2.8075917	0.5910720	1.6333966
Nearly best strategy	0.8734820	0.9800916	1.0025779	1.2803874	0.9730067	1.0692911

STANDARD DEVIATIONS

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5
Fixed order	0.0011192	0.0009078	0.0012641	0.0007833	0.0007833
Reverse fixed order	0.0006572	0.0006572	0.0006572	0.0008717	0.0014710
Random order	0.0014234	0.0011921	0.0015286	0.0011816	0.0017389
Increasing checking time	0.0011192	0.0009078	0.0012641	0.0007833	0.0007833
Nearly best strategy	0.0024537	0.0009083	0.0012662	0.0030003	0.0023371

RELATIVE STDs (Coeff. of variation)

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Averages
Fixed order	0.0948130	0.1389652	0.1136312	0.0530111	0.0530111	0.0906863
Reverse fixed order	0.0669196	0.0669196	0.0669196	0.1762222	0.0597813	0.0873525
Random order	0.1202132	0.1838517	0.1381589	0.2288664	0.0710856	0.1484351
Increasing checking time	0.0948130	0.1389652	0.1136312	0.0530111	0.0530111	0.0906863
Nearly best strategy	0.2387735	0.1390111	0.1136624	0.4452175	0.0960768	0.2065483

SQUARE ROOTS OF MSEs

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5
Fixed order	0.0011198	0.0009177	0.0012641	0.0095458	0.0102532
Reverse fixed order	0.0020526	0.0032212	0.0014486	0.0009274	0.0015227
Random order	0.0014255	0.0012060	0.0015293	0.0011859	0.0018201
Increasing checking time	0.0011198	0.0009177	0.0012641	0.0095458	0.0102532
Nearly best strategy	0.0028699	0.0009179	0.0012665	0.0033436	0.0024326

SQUARE ROOTS OF RELATIVE MSEs

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Averages
Fixed order	0.0951857	0.1376494	0.1137672	1.8137082	0.4101264	0.5140874
Reverse fixed order	0.1744722	0.4831824	0.1303754	0.1761979	0.0609081	0.2050272
Random order	0.1211650	0.1809013	0.1376341	0.2253149	0.0728036	0.1475638
Increasing checking time	0.0951857	0.1376494	0.1137672	1.8137082	0.4101264	0.5140874
Nearly best strategy	0.2439383	0.1376910	0.1139821	0.6352754	0.0973033	0.2456380

STRATEGIES' COMPARISON

AVERAGE VALUES (over life-tests)

STRATEGY TYPE	Total & masking	mean masking	1st risk fn. x 100	Total insp. time	2nd risk fn.	Total & wasted time	mean wasted time	%
Fixed order	751	1.50	0.49%	376.72	0.75	303.77	0.61	80.64%
Reverse fixed order	999	2.00	0.51%	398.63	0.80	272.96	0.55	68.47%
Random order	973	1.95	0.55%	364.32	0.73	275.78	0.55	75.70%
Increasing checking time	751	1.50	0.49%	334.15	0.67	261.21	0.52	78.17%
Nearly best strategy	659	1.32	0.29%	329.45	0.66	205.50	0.41	62.38%

LIFE TESTS PARAMETERS & STATISTICS

Simulation id. number:	2
Life test size:	500
Starting point for computation of the nearly best strategy:	25
Number of simulated life tests:	100
Average sum of system lifelengths:	8462.238

SYSTEM FEATURES

Number of system components:	5
System checking time limit:	0.95

COMPONENT FEATURES

Component	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$
Mean lifelength	85	150	90	190	40
Failure rate	0.01176	0.00667	0.01111	0.00526	0.02500
Inspection time	0.25	0.25	0.25	0.25	0.25

STATISTICS ON THE SIMULATED LIFELENGTHS

Component	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$
Average of sample means	85.07	150.53	90.42	189.51	40.22
Average of sample STDs	85.52	150.64	90.65	188.35	40.18

RESULTS ON THE FAILURE RATE ESTIMATORS

MEANS

ESTIMATORS / STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Average (for verification)
Fixed order	0.01189	0.00671	0.01068	0.01496	0.01496	0.01184
Reverse fixed order	0.00930	0.00930	0.01068	0.00514	0.02478	0.01184
Random order	0.01194	0.00674	0.01064	0.00504	0.02484	0.01184
Increasing checking time	0.01189	0.00671	0.01068	0.01496	0.01496	0.01184
Nearly best strategy	0.01185	0.00609	0.01059	0.00592	0.02476	0.01184

BIASES

ESTIMATORS / STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Averages (for verification)
Fixed order	0.0001220	0.0000468	-0.0004318	0.0096970	-0.0100398	-0.0001211
Reverse fixed order	-0.0024646	0.0026334	-0.0004318	-0.0001201	-0.0002226	-0.0001211
Random order	0.0001774	0.0000699	-0.0004706	-0.0002270	-0.0001554	-0.0001211
Increasing checking time	0.0001220	0.0000468	-0.0004318	0.0096970	-0.0100398	-0.0001211
Nearly best strategy	0.0000805	-0.0005785	-0.0005190	0.0006545	-0.0002431	-0.0001211

OBSERVED BIAS RATIOS E(failure rate' MLE)/(true failure rate)

ESTIMATORS / STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	(Geometric) Averages
Fixed order	1.0103713	1.0070267	0.9611421	2.8424385	0.5984081	1.6634053
Reverse fixed order	0.7905099	1.3950174	0.9611421	0.9771774	0.9910948	1.0265098
Random order	1.0150816	1.0104867	0.9576428	0.9568772	0.9937846	0.9340790
Increasing checking time	1.0103713	1.0070267	0.9611421	2.8424385	0.5984081	1.6634053
Nearly best strategy	1.0068388	0.9132244	0.9532939	1.1243491	0.9902752	0.9759359

STANDARD DEVIATIONS

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5
Fixed order	0.0011744	0.0009531	0.0012245	0.0008863	0.0008863
Reverse fixed order	0.0007534	0.0007534	0.0012245	0.0009295	0.0015352
Random order	0.0012675	0.0010512	0.0013124	0.0009918	0.0016958
Increasing checking time	0.0011744	0.0009531	0.0012245	0.0008863	0.0008863
Nearly best strategy	0.0012858	0.0016488	0.0013448	0.00162267 0.0015439	

RELATIVE STDs (Coeff. of variation)

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Averages
Fixed order	0.0988031	0.1419630	0.1146590	0.0592418	0.0592418	0.0947817
Reverse fixed order	0.0810083	0.0810083	0.1146590	0.1807387	0.0619584	0.1038745
Random order	0.1061371	0.1560493	0.1233434	0.1969280	0.0682560	0.1301427
Increasing checking time	0.0988031	0.1419630	0.1146590	0.0592418	0.0592418	0.0947817
Nearly best strategy	0.1085495	0.2708269	0.1269616	0.2741991	0.0623609	0.1685796

SQUARE ROOTS OF MSEs

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5
Fixed order	0.0011808	0.0009542	0.0012984	0.0097375	0.0100788
Reverse fixed order	0.0025772	0.0027391	0.0012984	0.0009373	0.0015513
Random order	0.0012799	0.0010536	0.0013943	0.0010174	0.0017030
Increasing checking time	0.0011808	0.0009542	0.0012984	0.0097375	0.0100788
Nearly best strategy	0.0012883	0.0017474	0.0014415	0.0017496	0.0015628

SQUARE ROOTS OF RELATIVE MSEs

ESTIMATORS STRATEGY TYPE	λ_1	λ_2	λ_3	λ_4	λ_5	Averages
Fixed order	0.1003655	0.1431316	0.1168523	1.8501178	0.4031537	0.5227242
Reverse fixed order	0.2190589	0.4108639	0.1168523	0.1780829	0.0620535	0.1973823
Random order	0.1087881	0.1580347	0.1254849	0.1933061	0.0681180	0.1307464
Increasing checking time	0.1003655	0.1431316	0.1168523	1.8501178	0.4031537	0.5227242
Nearly best strategy	0.1095049	0.2621069	0.1297307	0.3324285	0.0625115	0.1792565

STRATEGIES' COMPARISON

	AVERAGE VALUES (over life-tests)							
STRATEGY TYPE	Total & masking	mean	1st risk fn. x 100	Total insp. time	2nd risk fn.	Total & wasted time	mean	%
Fixed order	753	1.51	0.50%	310.63	0.62	248.81	0.50	80.10%
Reverse fixed order	657	1.31	0.33%	259.49	0.52	173.77	0.35	66.97%
Random order	700	1.40	0.41%	300.12	0.60	225.08	0.45	75.00%
Increasing checking time	753	1.51	0.50%	310.63	0.62	248.81	0.50	80.10%
Nearly best strategy	614	1.23	0.24%	249.73	0.50	153.12	0.31	61.31%

5. DISCUSSION

The simulation results indicate that, as far as economical aspects are concerned, it is undoubtedly the case to contemplate resorting to a best (or nearly-best) inspection strategy, since that is the way to obtain the most substantial savings in the inspection process. On the other hand, when precision only is concerned, while the best strategies do quite well, the best results are usually obtained with a random checking. Once this phenomenon has been observed, it is not difficult to find justifications of it. In fact, a random checking ensures a certain balance in the number of observed failures of each component, while other strategies which remain fixed with respect to the systems do not generally allow recording the failure times of the components systematically excluded from checking. Notice that even the best strategy is almost a fixed strategy in our particular simulation, for the well known characterizing property of the exponential distribution, see, for instance, Barlow and Proschan (1975). This is also the reason, indeed, why fixed order strategies tend to yield estimates having a lesser variance. In fact, in the absence of observed failure times for some components, it occurs that, for the components with no recorded failures, there are infinitely many solutions satisfying the ML equations, with the only constraint that they must add up to a given constant. Thus, it is only by the effect of the choice of equal starting values in the iterative algorithm used to solve the ML equations that we get identical values for the failure rate estimates of components with no recorded failures (see table with means of failure rate estimates). It is hence clear that the apparent lesser variability of the fixed type inspection strategies is mostly imputable to the absence of information about some failure rates. A side effect of that is also the lower bias generally obtained with the random strategy. Actually, the biases recorded with the random strategy are generally so much lower as to compensate for the higher variance and yield lesser mean square errors (see tables of biases and MSEs of failure rate estimates). On the economical side, however, the random checking performs quite inefficiently, in general; even though, it should be clear that there may exist several fixed strategies that perform even worse. The economic aspects should not be underestimated here, since, especially for high-tech equipment, the employment of highly specialized personnel in the inspection process may be very costly. The most noticeable advantage of an estimation improved by a planned search for the cause of failure is that, while it is extremely simple to apply, it may yield important savings in large-scale life-tests or in routine tests for quality control. Especially in automatized testing frameworks, it is always convenient to spend a few milliseconds of computing time to determine a convenient inspection strategy for each system, which may result in hundreds of hours of inspection time savings.

Finally, it turns out that the choice of the appropriate inspection strategy is largely a matter of the main concern of the experimenter. If economy is an issue, then a strategy of the type we have been referring to as best (or nearly-best) should be preferred. If the emphasis is on precision and costs are unimportant, then a "random" strategy can be used.

REFERENCES

Barlow, R.E., and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing - Probability models*, H.R.W., Inc., 58.

Basu, A., and Klein, J. (1982), Some recent development in competing risks theory, in *Survival Analysis*, eds. J. Crowley and R.A. Johnson, Hayward, CA: IMS, 216-229.

Berger, J.O., and Sun, D. (1993). Bayesian analysis for the poly-Weibull distribution, *Journal of the American Statistical Association,* **88**, 1412-1418.

Cox, D.R. (1959). The analysis of exponentially distributed lifetimes with two types of failures, *Journal of the Royal Statistical Society*, Ser. B, **21**, 411-421.

David, H.A. (1974). Parametric approaches to the theory of competing risks, *Reliability and Biometry - Statistical analysis of lifelengths,* edited by F. Proschan and R. J. Serfling, 275.

David, H.A., and Moeschberger, M.L. (1978), *The Theory of Competing Risks* (Griffin's Statistical Monographs & Courses No. 39), London: Charles W. Griffin.

Gastaldi, T., and Gupta, S.S. (1994). Minimax type procedures for nonparametric selection of the "best" population with partially classified data, *Center for Statistical Decision Sciences and Department of Statistics, Technical Report* #94-9C.

Gastaldi, T. (1993). Improved maximum likelihood estimation for component reliabilities with Miyakawa-Usher-Hodgson-Guess' estimators under censored search for the cause of failure, *Statistics & Probability Letters,* **19**, 5-18.

Guess, F.M., Usher, J.S., and Hodgson, T.J. (1991). Estimating system and component reliabilities under partial information on cause of failure, *J. Statist. Plann. Inference,* **29**, 75-85.

Mendenhall, W., and Hader, R.J. (1958). Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data, *Biometrika,* **45**, 504-520.

Miyakawa, M. (1984). Analysis of incomplete data in a competing risks model, *IEEE Trans. Reliability,* **33** (4), 293-296.

Usher, J.S., and Hodgson, T.J. (1988). Maximum likelihood analysis of component reliability using masked system life data, *IEEE Trans. Reliability,* **37** (5), 550-555.

Usher, J.S. (1993). On the problem of masked system life data, *Advances in Reliability*, Elsevier Science Publishers B.V., 435-443.

Usher, J.S., and Guess, F.M. (1989). An Iterative Approach for Estimating Component Reliability From Masked System Life Data, *Quality and Reliability Engineering International,* Vol. 5, pp 257-261.