Scand J Statist 12: 1–41, 1985

On Reliability Theory and its Applications[1]

BO BERGMAN

Linköping Institute of Technology

ABSTRACT. In this expository paper some results in reliability theory and their usefulness in industrial applications are discussed. First, some methods for the analysis of system reliability are reviewed. Then we discuss some methods applicable to the assessment of component reliability; Bayesian methods as well as stress–strength analyses and safety analysis of one-shot devices are discussed. Finally, we discuss some aspects on maintained systems including also a treatment of the modelling of software system reliability growth.

Key words: reliability theory, system reliability, component reliability, reliability review

0. Prologue: The industrial need

Man-made systems all suffer from imperfections, the origins of which may be manifold: the designer has neglected or has not been aware of some important facts concerning either the environment in which the system is to operate or the operation of the system itself; the manufacturer has introduced defects into the system when producing it; weaknesses are inherent in the materials from which the system is built. Often, these imperfections lead to the failure of the system to perform its function or they lead to hazards to its operators, users or others.

Another source of failure is natural component deterioration caused by, for example, friction and abrasive wear, metal fatigue or corrosion. Furthermore, failure may be caused be mistakes made by operators or maintenance personnel.

Very often, system or component failure is the result of many interacting factors. The failure of a dynamically loaded component may, for example, be a result of a defect being introduced during production, leading to crack initiation, a growing fatigue crack accelerated by a corrosive environment and, finally, an extreme load exceeding the residual strength of the component.

Most failures have economic consequences, not only because of the necessary repair or component replacement following a failure but, more important, because of other failure effects such as a loss of production. As a rule of thumb, maintenance expenditure for machinery and equipment amounts to about 50% of total investment costs. According to a recently performed investigation by the Danish Maintenance Association (DDV), maintenance expenditure accounts for about 6–8% of the Gross National Product (GNP). But, probably, this expenditure is small in comparison with the cost of losses of production due to failures.

Today, many buyers are aware of the amounts payed for maintenance and, consequently, they use not only acquisition costs but total life cycle costs (LCC), including maintenance costs, for the evaluation of industrial proposals. This routine for purchase is used, for instance, by two large buyers in Sweden—the Swedish Defence Material Administration and the Swedish State Railway. Also producers have found that they become more competitive by using the LCC concept.

---

[1]This paper is an updated version of an invited lecture at the 10th Nordic Conference on Mathematical Statistics, Bolkesjø, Norway, 19–21 June 1984. The lectures were concluded with the discussion which is reported at the end of the paper.

For many complex systems such as nuclear power plants, off-shore systems, aircraft, and medical instrumentation, the consequences of failures are not just economic in nature but they may also affect the safety of human beings. Also the failure of comparatively minor systems and products can have serious consequences; even an electric razor may be dangerous if, for example, a failure occurs in the electrical insulation.

Safety issues have, during the last few decades, become increasingly important. Safety has become a major topic of public opinion, partly as a result of the debate concerning nuclear power but also because of accidents which have occurred to less controversial systems. Governmental requirements for safety analyses of systems are increasing in importance and in Europe legislation seems to be moving towards a strict product liability, similar to the situation in the USA, where a manufacturer whose product causes injury to a user can anticipate large compensation claims.

These facts constitute an incentive and a need for industry to perform systematic studies for the identification and possible elimination of failure causes, quantification of failure occurrences and for the reduction of failure consequences. These systematic studies should be performed parallel to the design process in order to affect the final design in the most economic way. This industrial need has brought about the creation of the reliability engineer, an expert on the study of failures, their causes, effects, quantification and elimination. As one of the tools of the reliability engineer, reliability theory has evolved. It is the aim of this exposition to discuss some topics in reliability theory and their applications in industry. It should be emphasized that this discussion is by no means exhaustive. Rather, it reflects some of the interests of the author.

1. Introduction

As indicated in the prologue, reliability analysis has to be performed in parallel with the design rocess in order to support important decisions. Thus, at the beginning of a project, the nalysis has to be performed with only a small amount of information available. On the basis of perational requirements, the reliability analyst should recommend a reliability structure (for example, redundancies) and reliability targets for subsystems.

As the design process proceeds, the reliability engineer is able to do more and more detailed analysis. Essentially, this may be described as four different activities: (1) find those events at system level which are important from a safety or economic point of view and evaluate their consequencies; (2) find those (sequences of) events (or failures) at component level which cause these serious system events; (3) find reliability figures for these component events either by using experiences from similar components in operation or test or from some sort of stress–strength analysis; (4) evaluate the proposed design and, if feasible, recommend redesign or other changes.

In general, reliability theory may help the reliability engineer carry out the three latter types of activities: the reliability structure of the system with respect to different system events may be found by performing fault tree analyses and the critical sequence of events may be found (see section 2). The systematic analysis of experiences from systems and components in operation or of reliability tests is supported by statistical reliability theory as indicated in section 3. In section 4 we discuss some topics in reliability theory useful for cost effectiveness studies of maintained systems, i.e. topics useful for the evaluations and recommendations as under the fourth point above.

Throughout, we use increasing in place of non-decreasing, decreasing instead of non-increasing, positive in place of non-negative and negative in place of non-positive.

2. System reliability

Let us study the reliability of a technical system. Assume that *at any given instant* the system is either functioning or faulty. Denote the state of the system by ϕ and set

$$\phi = \begin{cases} 1, & \text{if the system is functioning;} \\ 0, & \text{if the system is faulty.} \end{cases}$$

Also, assume that the system may be decomposed into components numbered as $1, 2, \ldots, r$, and that at any given instant each component is either functioning or faulty. Let x_i denote the state of the ith component:

$$x_i = \begin{cases} 1, & \text{if component number } i \text{ is functioning;} \\ 0, & \text{if component number } i \text{ is faulty.} \end{cases}$$

We write $\mathbf{x}=(x_1, \ldots, x_r)$, the vector of component states. Furthermore, assume that the state of the system is completely determined by the states of its components: $\phi=\phi(x_1, \ldots, x_r)$. The binary function ϕ: $\{0, 1\}^n \to \{0, 1\}$ is called the (reliability) structure function of the system. The two simplest types of systems are the series and parallel systems for which we have

$$\phi(\mathbf{x}) = \prod_{i=1}^{r} x_i = \min_i x_i,$$

and

$$\phi(\mathbf{x}) = 1 - \prod_{i=1}^{r} (1-x_i) = \max_i x_i,$$

respectively. The latter function is often written

$$\coprod_{i=1}^{r} x_i = 1 - \prod_{i=1}^{r} (1-x_i);$$

here the symbol $\coprod$ is read "ip". For ϕ to be a structure function of a real system we require $\phi(\mathbf{0})=0$ and $\phi(\mathbf{1})=1$; here $\mathbf{0}=(0, \ldots, 0)$ and $\mathbf{1}=(1, \ldots, 1)$. Very often it is also assumed that

(i) each component is relevant, i.e. for any i, $i=1, \ldots, r$ an $\mathbf{x}$ exists such that

$$\phi(1_i, \mathbf{x}) - \phi(0_i, \mathbf{x}) = 1,$$

where $(1_i, \mathbf{x})=(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_r)$ and $(0_i, \mathbf{x})$ is defined correspondingly, and

(ii) that for any state vectors $\mathbf{x}_1 \geq \mathbf{x}_2$ we have $\phi(\mathbf{x}_1) \geq \phi(\mathbf{x}_2)$; here $\mathbf{x}_1 \geq \mathbf{x}_2$ means coordinatewise inequality: $x_{11} \geq x_{21}, \ldots, x_{1r} \geq x_{2r}$.

Systems fulfilling (ii) are called *monotone* and monotone systems fulfilling (i) are called *coherent*.

A theory for coherent systems was first given by Birnbaum *et al.* (1961). Nice presentations are given in e.g. Barlow & Proschan (1981b) and Kaufmann *et al.* (1977). Below we shall give a short introduction; for a more complete description we refer to the above mentioned books.

To be precise, we define a system as (C, ϕ), where $C=\{1, \ldots, r\}$ denotes the set of (indices of) components and ϕ the structure function. For any state vector $\mathbf{x}$ we define $C_0(\mathbf{x})=\{i \in C \mid x_i=0\}$ and $C_1(\mathbf{x})=\{i \in C \mid x_i=1\}$. A *cut vector* is an $\mathbf{x}$ such that $\phi(\mathbf{x})=0$ and $C_0(\mathbf{x})$ is the corresponding *cut set*. Similarly, we define a *path vector* $\mathbf{x}$ such that $\phi(\mathbf{x})=1$ with the *path*

set $C_1(\mathbf{x})$. A *minimal cut vector* $\mathbf{x}$ is such that for any $\mathbf{y}>\mathbf{x}$ (meaning $y_1\geq x_1,\ldots,y_r\geq x_r$, with $y_i>x_i$ for some i, $1\leq i\leq r$), we have $\phi(\mathbf{y})=1$; we call $C_0(\mathbf{x})$ a *minimal cut set*. A *minimal path set* is defined correspondingly. Let $\mathscr{K}_1,\ldots,\mathscr{K}_k$ be the minimal cut sets of $\{C,\phi\}$ and let $\mathscr{P}_1,\ldots,\mathscr{P}_\varrho$ be the minimal path sets. We can now represent the structure function ϕ as

$$\phi(\mathbf{x})=\prod_{j=1}^{k}\coprod_{i\in\mathscr{K}_j} x_i=\coprod_{j=1}^{\varrho}\prod_{i\in\mathscr{P}_j} x_i;$$

Often it is possible to decompose a system (C,ϕ) into a number of modules, i.e. disjoint subsystems (A_i,χ_i), $i=1,\ldots,m$, and an organizing structure ψ, such that $\cup_{i=1}^{m}A_i=C$, and $\phi(\mathbf{x})=\psi\{\chi_1(\mathbf{x}^{A_1}),\ldots,\chi_m(\mathbf{x}^{A_m})\}$; here $\mathbf{x}^{A_i}$ is the vector of component states for the components in A_i. Obviously such a decomposition is valuable when analysing a system.

The theory of coherent systems are closely related to network theory. A reliability network defined on a set C of components consists of a set of vertices among which two are distinguished and called source (s) and terminal (t), and a set of a arcs between vertices such that each arc corresponds to one component. A component, however, may be associated with many arcs. The reliability network is functioning if a path of functioning arcs between s and t exists—an arc is functioning if the corresponding component is functioning. In this case it is possible to establish an equivalence between coherent structures and reliability networks (see e.g. Kaufmann *et al.*, 1977). Often, however, independence between the states of the arcs is assumed. In this case we have exactly one arc corresponding to a given component. These networks are much less general than coherent systems. For example, to represent a k-out-of-n-system (a system which is functioning if at least k of its n components are functioning) as a reliability network it is necessary to have more than one arc representing the same component.

Remark. Also reliability networks in which more vertices than s and t are distinguished may be of interest. In a computer network we may be interested in the event that *all* computers do communicate. In a *K-terminal* network a set K of vertices are distinguished and the network is functioning if between any two vertices in K it is possible to find a path of functioning arcs (see e.g. Satyanarayana & Chang, 1983). □

Obviously, binary coherent systems are very important in applications, but they are not the only ones deserving attention. They have some serious drawbacks: only binary state vectors are allowed and the systems are given an essentially static description, i.e. the time order of failure events are not easily taken into account.

A component or a system may be multistate either because it is degenerating due to e.g. wear or because it may take different failure states as e.g. a relay which has two failure states, "short-circuit" and "open". The first type of multistate systems has been studied by e.g. Barlow & Wu (1978) and El-Neweihi *et al.* (1978). Natvig (1982) has given a general definition of multistate coherent systems with ordered component and system states (see also Natvig, 1984).

Calderola (1980) has studied systems with non-ordered component states as in the case of relays. To each component state a Boolean variable is associated. Since a component is assumed to take at most one failure state a restriction is imposed upon these Boolean variables. Also Calderola give a generalization of the coherent system concept.

Since most technical systems are multistate in one meaning or another these generalizations are very valuable. Perhaps it should be even more natural from an applied point of view to assume a partial ordering on the set of (component) states taking into account both of the discussed types of multistate behaviour.

The other problem concerning the coherent system description mentioned above is its

essentially static approach. Non-relevant components may e.g. be very important to the system if they are *protective*, i.e. they protect other components from an otherwise severe environment. Examples in electronic systems may e.g. be diodes which have been built into a system just to take away transients. If such a diode has failed the system may still be functioning after the failure if it was functioning just before, *but* the future reliability prognosis of the system has become much worse. Other types of devices protecting components from overcurrent, overvoltage, vibrations and corrosion may be found in many electrical and mechanical systems. Thus also non-relevant components (in the meaning of coherent system theory) has to be taken into account, when time properties of systems are studied. Some systems change their characteristics in such a way that the time order of the failure events become important. Such systems have become quite common with the introduction of microcomputers; important examples occur in avionics. It is not obvious how this dynamic structure should be taken into account in a general system reliability theory.

Now, let us for a moment concentrate on the question how to find the reliability structure of a system. This question has got an excellent answer in the so-called Fault Tree Analysis (FTA) technique—a technique for building a logic tree relating an undesired system event to sequences of contributing events. This is performed by making successive logical decompositions of the higher system level events into contributing (lower system level) events using primarily "and" and "or" operators until eventually component failure events are reached. Each step in the decomposition should be as small as possible in order to reduce the risk that a contributing sequence of events is forgotten. General guidelines for the development of a fault tree are given by Haasl *et al.* (1981). Today, the FTA has become one of the most important tools for the reliability analyst.

It should be noted that in the system description obtained by the FTA different basic events may concern the same component possibly in different failure modes. Also complementary events may occur. However, in the following we restrict to the special case when the FTA in fact gives a description of a coherent system, which is the case if the components have only one failure state and no complementary events occur. In this case the failure situations of the system may be represented, e.g. by the minimum cut sets of the system. These may be obtained by using some suitable algorithm (see e.g. Fussel & Veseley, 1972). However, the problem of finding all the cut sets of a fault tree has been shown by Rosenthal (1975) to be NP difficult, i.e. no general algorithm for finding the cut sets of a fault tree exists such that its running time is bounded by some polynomial function of the number of basic events in the system. Another problem concerning the cut sets is their number; also fault trees moderate in size may have an incredible number of cut sets. The reason we still find the cut set representation valuable will be obvious later.

Before considering the quantitative evaluation of a fault tree we shall discuss one of the most problematic questions in FTA—that of dependency. Esary *et al.* (1967) introduced the concept of (positive) association; random variables (r.v.'s) $T_1, \ldots, T_n$ are (positive) *associated* if for any pair of increasing binary functions Γ and Δ we have

$$\text{Cov}\,(\Gamma(\mathbf{T}), \Delta(\mathbf{T})) \geqslant 0;$$

here $\mathbf{T}=(T_1, \ldots, T_n)$. Verifying association using this definition may be a hard task. However, a sufficient criterion, which is very useful, is the following: if r.v.'s $T_1, \ldots, T_n$ are conditionally increasing in sequence, i.e. if T_i is stochastically increasing in $T_1, \ldots, T_{i-1}$, then $T_1, \ldots, T_n$ are associated (see further, Barlow & Proschan, 1981b). Obviously association is some sort of positive dependence; associated r.v.'s tend to behave similarly as for example minimum cut structures of a coherent system having components in common and the failure of components subjected to the same stress environment.

Remark. Negative association was introduced and studied by Joag-Dev & Proschan (1983). We shall not use that concept here. In the following association means positive association.

In a way (positive) association is the natural dependence concept in reliability theory. Let X_1 and X_2 be two binary r.v.'s; they are associated if and only if they are positively correlated, which means that

$$P(X_1X_2=1)\geqslant P(X_1=1)\ P(X_2=1).$$

More generally, for a series structure of associated components (or subsystems) it is conservative to assume independence, i.e. the failure probability is overestimated. This means that in order to obtain a conservative (lower) bound for system reliability when component states are associated r.v.'s, it is sufficient to model the dependence between components within the same cut sets, since

$$\prod_{j=1}^{k} P\left(\coprod_{i=\mathscr{K}_j} X_i=1\right)\leqslant P\{\phi(\mathbf{X})=1\};$$

here we have used that increasing functions of associated r.v.'s are associated. □

The analyst has to find and separate out those common factors which may affect the reliability of more than one component in the same cut set. If for example two components may fail due to the same environmental shock we have to define a new basic event "environmental shock fatal to both components". Now, hopefully, we can regard the basic events independent. Otherwise, we have to find and model further common failure causes. The identification and modelling of possible dependencies between failure events is usually the most crucial step in an FTA, and it is not always possible to separate out the common factors in the way indicated above; special models have to be used.

Using the same arguments as above we can also find an upper bound for the system reliability, i.e.

$$\prod_{j=1}^{k} P\left(\coprod_{i\in\mathscr{K}_j} X_i=1\right)\leqslant P\{\phi(\mathbf{X})=1\}\leqslant \coprod_{j=1}^{\varrho} P\left(\prod_{i\in\mathscr{P}_j} X_i=1\right),$$

which reduces to

$$\prod_{j=1}^{k}\coprod_{i\in\mathscr{K}_j} p_i\leqslant P\{\phi(\mathbf{X})=1\}\leqslant \coprod_{j=1}^{\varrho}\prod_{i\in\mathscr{P}_j} p_i,$$

if $X_1,\ldots,X_r$ are independent with $E(X_i)=p_i$. These inequalities may be improved, e.g. by using modular decompositions and some min–max bounds utilizing that for any $\mathbf{x}$

$$\min_{i\in\mathscr{P}_j} x_i\leqslant\phi(\mathbf{x})\leqslant \max_{i\in\mathscr{K}_{j'}} x_i,$$

for all j, j' (see Natvig, 1980).

We can also use the inclusion–exclusion method and Bonferroni's inequalities to obtain upper and lower bounds for system reliability.

When the random component state indicators X_i, $i=1,\ldots,r$, are independent we define $h(\mathbf{p})=E\{\phi(\mathbf{X})\}=P\{\phi(\mathbf{X})=1\}$; here $\mathbf{p}=(p_1,\ldots,p_r)$ and $p_i=P(X_i=1)$, $i=1,\ldots,r$.

In order to make *exact* reliability computations for coherent systems the *factoring algorithm* seems to be the most appropriate. The factoring algorithm is based upon the fact that we can always write

$$\phi(\mathbf{X})=X_i\phi(1_i,\mathbf{X})+(1-X_i)\,\phi(0_i,\mathbf{X}),$$

and thus for independent component state indicators we have

$$h(\mathbf{p})=p_ih(1_i,\mathbf{p})+(1-p_i)\,h(0_i,\mathbf{p}).$$

This is often called a *pivotal decomposition*, and the component i is called a *pivotal element*.

Further, the algorithm utilizes series and parallel reductions, i.e. modules with series or parallel structure are identified and replaced by "super"-components. These reductions are applied first on $\phi(\mathbf{x})$ itself and, after the pivotal composition, on $\phi(1_i, \mathbf{x})$ and $\phi(0_i, \mathbf{x})$. Then further pivotal decompositions are performed and so on until single components or "super"-components are obtained (see Fig. 1). In a graph theoretical context Satyanarayana & Chang (1983) have shown that, in a K-terminal network with undirectional and independent arcs, successive pivotal elements may be chosen such that the number of operations is of size $O(b+r)D$, where r is the number of arcs, b is the number of vertices, D is a complexity measure of the graph called its *domination*, and, as usual, $O(\cdot)$ is a function such that $O(x)/x \to \text{const.}$ when $x \to \infty$.

The domination D is equal to the absolute difference between the numbers of odd and even *formations* of the network, where a formation is a set of minimal paths $\mathscr{P}_{i_1}, \ldots, \mathscr{P}_{i_f}$ such that $\cup_{l=1}^{f} \mathscr{P}_{i_l} = C$; a formation is odd (even) if f, the number of path sets in the formation, is odd (even).

Generally, it can be shown (see Huseby, 1983), that if all the substructures obtained by the successive pivotal decompositions of the system are coherent, then D is equal to the number of "leaves" of the tree corresponding to the algorithm, and the algorithm is optimal. Very recently, Huseby (1984) has shown by using matroid theory (see e.g. Welsh, 1976) that optimal algorithms may be found for a large class of coherent systems including K-terminal networks and "k-out-of-n" structures. Optimal algorithms for the latter class of structures have been given by Wu (see Barlow, 1982).

Now, let us return to the design process. In order to support the designer the reliability analyst shall not only present reliability figures but often he also has to suggest where to allocate resources for reliability improvements, e.g. give a suggestion on which components to improve from a reliability point of view. To do this he needs a measure on the relative importance of each component with respect to system reliability.

Birnbaum (1969) suggested two importance measures, one purely structural and the other probabilistic: with respect to a certain state vector $(1_i, \mathbf{x})$ the component i is *critical* if $\phi(1_i, \mathbf{x}) - \phi(0_i, \mathbf{x}) = 1$. For each component we can determine 2^{r-1} different state vectors $(1_i, \mathbf{x})$ and the proportion of these for which the state i is critical is a natural indicator on the *structural importance* of the component i. However, the failure probabilities of the

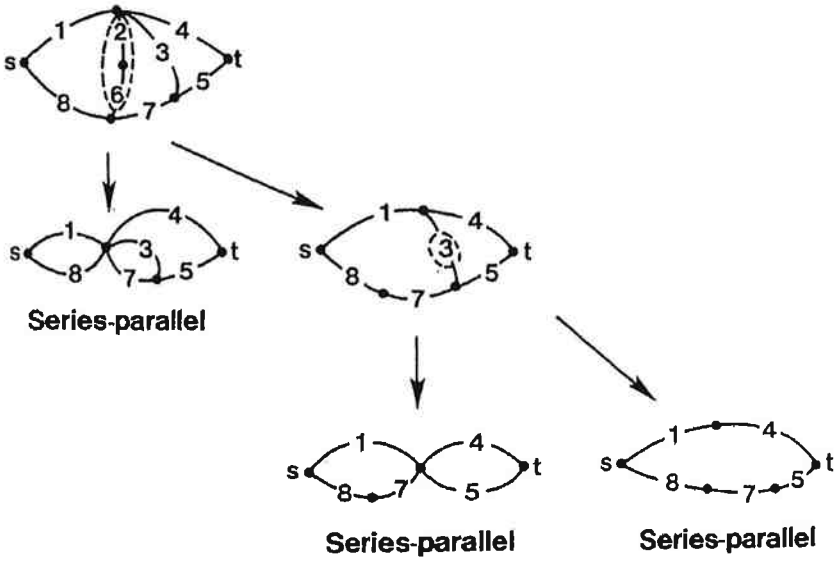

Fig. 1. Illustration of the factoring algorithm.

components are not taken into account. Birnbaum (1969) therefore suggested the following (probabilistic) importance measure of the ith component:

$$I_B^{(i)}=P\{\phi(1_i, \mathbf{X})-\phi(0_i, \mathbf{X})=1\},$$

i.e. the probability that the ith component is critical with respect to the random state vector $\mathbf{X}$. Usually independence between failure events are assumed and then it is easily seen that

$$I_B^{(i)}=\frac{\partial h(\mathbf{p})}{\partial p_i}.$$

But, a component may contribute to a system failure event without being critical. Veseley & Fussel (see e.g. Fussel, 1975), suggested the following importance measure:

$$I_{VF}^{(i)}=P\{i\in C_0(\mathbf{X})\,|\,\phi(\mathbf{X})=0\},$$

which takes this fact into account. However, this measure gives the same importance to all the components of a parallel system irrespective of failure probability.

Until now we have studied the system at a given point t in time, but that is not quite relevant for most design or redesign decisions; a time-independent importance measure would be more suitable. In the following we shall give explicit reference to the time t. Assume that each component has (random) life τ_i with cumulative distribution function (c.d.f.) F_i and that $X_i(t)=1_{\{\tau_i>t\}}$, the indicator variable of the event $\{\tau_i\geq t\}$. Barlow & Proschan (1975) suggested the following time-independent importance measure

$$I_{BP}^{(i)}=\int_0^\infty I_B^{(i)}(t)\,dF_i(t);$$

here $I_B^{(i)}(t)$ is the Birnbaum reliability importance measure with an explicit reference to time t. We have

$$I_{BP}^{(i)}(t)=\int_0^\infty (h\{1_i, \bar{\mathbf{F}}(t)\}-h\{0_i, \bar{\mathbf{F}}(t)\})dF_i(t)$$

$$=P\text{ (the failure of the } i\text{th component coincides with the failure of the system);}$$

here $\bar{\mathbf{F}}(t)=\{\bar{F}_1(t), \ldots, \bar{F}_r(t)\}$. Since $I_{BP}^{(i)}(t)$ is based on $I_B^{(i)}$ it has the same type of disadvantage; it does not take into account that the failure of the ith component may contribute to the failure of the system even if it does not coincide with the failure of the system. A time independent importance measure having characteristics similar to the time-dependent Veseley–Fussel measure has been proposed by Natvig (1979): let Z_i be the reduction in remaining system life due to the failure of the ith component. This may be interpreted in the following way: let $\tau_{(i,2)}$ be the time to the *second* failure of the ith component if the first failure is immediately repaired to the condition just before the first failure. Then,

$$Z_i=\tau_\phi^{(i)}-\tau_\phi,$$

where τ_ϕ and $\tau_\phi^{(i)}$ are the life lengths of the system based on component lifes $(\tau_1, \ldots, \tau_r)$ and $(\tau_1, \ldots, \tau_{i-1}, \tau_{(i,2)}, \tau_{(i+1)}, \ldots, \tau_r)$ respectively. Natvig suggests as an importance measure:

$$I_N^{(i)}=E(Z_i)\Big/\sum_{j=1}^n E(Z_j).$$

This measure has the advantage that also the importance of e.g. protective and (in the meaning of coherent system theory) non-relevant components may be calculated and all types of contribution by the failure of the ith component on the system failure affect the importance

measure. However, also reductions other than Z_i may be natural (see e.g. Natvig, 1982). An alternative measure of the reduction due to the failure of the ith component is given by $Y_i=\tau_{\phi,i}-\tau_\phi$, where τ_ϕ is defined as above and $\tau_{\phi,i}$ is the life of the system if the components have lives $\tau_1,\ldots,\tau_{i-1},\infty,\tau_{i+1},\ldots,\tau_r$, i.e. $\tau_{\phi,i}$ is the life of the system if the ith component is replaced by a perfect one.

We can conclude that reliability importance measures are important to the reliability analyst but that some questions on the most suitable one remain to be answered.

3. Component reliability

For the purpose of modelling component failure events we can distinguish at least three important situations corresponding to (i) one-shot devices, (ii) non-repairable devices with non-zero time to failure, and (iii) repairable devices. In section 4 we shall discuss reliability characteristics of repairable units and in this section we shall discuss the two other situations.

3.1. Life distributions

The theory of life lengths of components is well developed (see e.g. Barlow & Proschan, 1981b); here we shall only give some basic concepts and some comments with respect to applications. Let τ be the random time to failure of a unit under study, with a corresponding *life distribution* F (i.e. $F(0-)=0$), and assume that F is absolutely continuous (a.c.) with probability density (p.d.)f. One of the most important concepts in reliability theory is the *failure rate function* $r(t)=f(t)/\bar{F}(t)$, $t\geq 0$; here $\bar{F}(t)=1-F(t)$, $t\geq 0$, is the *survival function* of τ. Since

$$P\{\tau\in(t,\ t+\Delta t)\,|\,\tau>t\}=r(t)\ \Delta t+o(\Delta t),$$

where as usual $o(\Delta t)/\Delta t\to 0$ as $\Delta t\to 0$ we may interpret $r(t)$ as a natural indicator of the ageing of the unit, or, in other words, "its proneness to failure". Some typical cases are given below.

(a) Constant failure rate modelling the life of a unit which is not affected by age, i.e. the life distribution of τ, is the exponential cdf $F(t)=1-\exp(-\lambda t)$, $t\geq 0$; here λ is the *failure rate* and $r(t)=\lambda$, $t\geq 0$. This is the family of life distributions most used in applications.

(b) Decreasing Failure Rate (DFR); a life distribution (not necessarily a.c.) is said to be DFR if $\bar{F}(t+x)/\bar{F}(t)$ is increasing in $t\geq 0$ for each $x\geq 0$. Often electronic devices are assumed to have DFR life distributions. This may be interpreted as a consequence of a varying manufacturing quality; low quality devices fail early and, after a while, only high quality devices are left in the population. Note that a mixture of e.g. exponential distributions with different failure rates is DFR (see Proschan, 1963).

(c) Increasing Failure Rate (IFR); a life distribution (not necessarily a.c.) is said to be IFR if $\bar{F}(t+x)/\bar{F}(t)$ is decreasing in $t\geq 0$ for each $x\leq 0$. This is a life length model for devices ageing because of, for example, wear, fatigue or cumulative damage.

(d) Bath-tub-shaped Failure Rate (BFR); a life distribution is BFR if it is a.c. on $0<t<x_2$ and if $r(t)$ is decreasing on $0<t<x_1$ and increasing on $x_1\leq t\leq x_2$ for some x_1, x_2; $0\leq x_1\leq x_2\leq\infty$, such that $\bar{F}(x_2)=0$. This is the most referred to non-parametric class of life distributions in reliability engineering texts. In fact, any IFR or DFR life distribution belongs to the BFR class.

(e) Increasing Failure Rate Average (IFRA); a not necessarily a.c. life distribution is IFRA if $\bar{F}(t)^{1/t}$ is increasing in $t\geq 0$. For an a.c. life distribution this means that $\int_0^t r(s)\,ds/t$ is increasing in $t\geq 0$. It is perhaps somewhat astonishing that a life distribution of a coherent system built of components with IFR life distribution is not necessarily IFR.

For instance, a redundant system consisting of two components with independent exponential life lengths with different parameters is not IFR. However, the IFRA class is closed under formation of coherent systems (see e.g. Ross, 1972).

Other non-parametric classes of interest in special applications are e.g. DMRL (Decreasing Mean Residual Life: $e(t) \searrow$ as $t \nearrow$; here $e(t)$ is the mean residual life, i.e. $e(t)=\int_t^\infty \bar{F}(s)\,ds/\bar{F}(t)$), NBU (New Better than Used: $\bar{F}(t) \geq \bar{F}(t+s)/\bar{F}(s)$, for any $t \geq 0$ and $s \geq 0$), HNBUE (Harmonic New Better than Used in Expectation: $\int_t^\infty \bar{F}(s)\,ds \leq \mu \exp(-t/\mu)$; here $\mu=e(0)$), and natural duals to these classes. For a discussion of characterizations and closure properties of these classes see e.g. Barlow & Proschan (1981b) and Haines & Singpurwalla (1974). For some further references see also Bergman & Klefsjö (1984b). The parametric families most used in reliability applications are the exponential, Weibull, Gumble (i.e. extreme value type I) lognormal, normal and gamma distributions. The most relevant arguments for the exponential life distribution emanates from limiting properties of superposing renewal processes (a theorem due to Grigelioni; see e.g. Barlow and Proschan, 1981b), thinning renewal processes (see e.g. Råde, 1972), or simply the adequacy of the exponential model for rare coincidencies (see e.g. Keilson, 1979). The Weibull and Gumble distributions are extreme value distributions and thus natural life distributions wherever minima are considered. The Central Limit Theorem gives argument for the use of the lognormal and normal distributions.

Remark. To be correct, the normal and Gumble distributions are not life distributions, but they may give useful approximations. □

Very often, however, theoretical arguments seem appropriate, but the population under study is not homogeneous, i.e. the theoretical arguments apply to subpopulations and the appropriate life distribution should in fact be a mixture (see e.g. Hahn and Meeker, 1982). But sometimes mixtures are used in applications even if they are not appropriate. For instance, in populations of electronic components often early failures are observed due to quality defects. Often the life distribution of such a population is modelled as a mixture of e.g. two Weibull distributions, one modelling the subpopulation of substandard quality and the other the main population (see e.g. Hahn & Meeker, 1982; Jensen & Petersen, 1982). However, this model has to be handled with great care, since the *right* tail of the mixture distribution may very well be determined by the distribution modelling the *substandard* quality subpopulation. Let $W_i(t)=1-\exp\{-(t/\alpha_i)^{\beta_i}\}$, $i=1, 2$, be the two Weibull distributions in the mixture and assume e.g. that $\alpha_1<\alpha_2$ and $\beta_1<\beta_2$. Then W_1 corresponds to the substandard quality but it is easily seen that the extreme *right* tail of the mixture distribution is determined by W_1, i.e. the mixture distribution has an extreme right tail which is longer than that of the main population. This pecularity is perhaps not too serious if the model is used only to determine a procedure for eliminating the substandard subpopulation from the main population, so-called burn-in (see section 3.4). But, if this model is used in cases when also the right tail is important, serious mistakes can be made.

A competing risk model would be much more appropriate: let τ_1 be the time to failure of a component *without* quality defects and τ_2 the time to failure with respect to a quality defect. For a component without quality defects $\tau_2=\infty$. Then the life of a random component is equal to $\tau=\min(\tau_1, \tau_2)$. It is natural to assume that τ_1 and τ_2 are independent.

However, as noted also by Lawless (1983), more emphasis should be put on models taking into account early failure mechanisms. Other models not studied enough for reliability applications are competing risk models for components with many failures modes, and regression models, e.g. of the type suggested by Cox (1972), are probably very useful for the study of accelerated life tests.

As mentioned already in section 2, the modelling of dependencies is very important in reliability applications. Therefore multivariate life distributions are of great concern. During the last decade interesting theoretical results of multivariate generalizations of the natural non-parametric classes of life distributions have been obtained (see e.g. Block & Savits, 1981). However, it seems the direction of generalizations are so many that to each univariate class it corresponds a lot of multivariate classes. This multitude makes that approach less attractive. It seems as if one of the most promising approaches is that of Marshall & Olkin (1967), who introduced the bivariate exponential corresponding to a certain shock model: two components are exposed to three types of shocks accruing according to three independent Poisson processes with intensities λ_1, λ_2 and λ_{12}, respectively. At the ith type of shock the ith component fails, $i=1, 2$, while if the third type of shock occurs both components fail. This model easily extends to more than two components. Observe that this model is closely related to the procedure discussed in section 2 for modelling dependencies in a fault tree.

Another multivariate model is discussed in section 3.3.

3.2. Statistical inference on life distributions

In a recent paper Lawless (1983) reviews advances made during the last 25 years with respect to statistical methods in reliability. He also suggests some areas where further work is needed. Here we shall only discuss some subject matters not much penetrated in that paper.

In most reliability applications we need component reliability information in order to support some kind of decision concerning e.g. design, redesign or maintenance planning. Often we have information not only from current data but also from more or less appropriate experiences and handbook suggestions. Therefore, it is natural to take a Bayesian point of view in reliability. A compendium of Bayesian methods in reliability has recently been provided by Martz & Waller (1982). A thorough treatment of some basic concepts in Bayesian inference on reliability is given by Barlow & Proschan (1980, 1981a).

Before discussing different inference techniques we have to choose a model. This choice may be supported by theoretical arguments and by data, preferably both ways. While choosing a model it is very important to keep in mind what sort of question (i.e. what sort of decision problem) we want to answer by using the model, the data, and other pieces of information at our disposal. This seems to be a trivial comment, but very often it is forgotten in many applications. If e.g. the extreme left tail of the life distribution is of importance to us because of a safety issue we should not choose a model such that properties of the right tail affect the result of our analysis; an overdue belief in a statistical model might give very misleading results. Some pertinent examples given by Österberg & Öfverbeck (1977) concerning strength distributions are worth considering.

In order to choose a model fitting the data or to evaluate a model against the data, different techniques may be appropriate; often graphical methods are to be preferred. Three graphical methods have proven to be very valuable: probability plotting, specially when tail properties and percentiles are considered, Total Time on Test (TTT) plotting, when considerations on the shape of the failure rate function are of interest, and hazard plotting, for identification purposes when data are censored.

The first two techniques can be said to be problem oriented, while the latter provides a very simple method for evaluating censored data; it is closely related to the point process approach to censored survival data analysis (see e.g. Aalen, 1978). Probability plotting is well known and hazard plotting suggested by Nelson (1972) is presented in Nelson (1982). Here we shall only present some illustrative results concerning the use of the TTT-plotting technique originally suggested by Barlow & Campo (1975). In section 3.4 we shall give a new example

(see also Bergman & Klefsjö, 1984c) illustrating its use on burn-in problems, i.e. problems concerning the removal of early failures, applicable to electronic devices.

Let $t_{(1)}, \ldots, t_{(n)}$ be an ordered sample from a life distribution F and let $t_{(0)}=0$, then the *total time on test* to the ith failure, is equal to

$$T_i = \sum_{j=1}^{i} (t_{(j)} - t_{(j-1)})(n-j+1), \qquad i=1, \ldots, n, \; T_0=0.$$

Assuming that F is the exponential distribution we find that $u_i = T_i/T_n$, $i=1, \ldots, n-1$, are distributed as an ordered sample from a uniform distribution on (0, 1). Thus, plotting u_i against i/n, $i=0, \ldots, n$, we expect the plotted positions to be close to the diagonal in the unit square. A TTT plot is obtained by connecting the plotted positions with straight lines. Generally, we can write

$$u_i = \int_0^{F_n^{-1}(i/n)} \{1-F_n(s)\}\, ds/\bar{t};$$

here $\bar{t}=\sum_{i=1}^{n} t_i/n$, F_n is the empirical distribution function and for any c.d.f. G we define $G^{-1}(u)=\inf\{t;\, G(t)\geq u\}$.

For a general F and for n increasing it is possible to prove that the TTT plot approaches

$$\phi(u) = \int_0^{F^{-1}(u)} \{1-F(s)\}\, ds/\mu, \qquad 0\leq u\leq 1,$$

uniformly with probability 1 (wp1) (see Barlow *et al.*, 1972, Ch. 5); here $\mu=\int_0^\infty \{1-F(t)\}\, dt$, the expectation of the r.v. τ with c.d.f. F, and $\phi(\cdot)$ is called the *scaled TTT transform* of F. This fact may be used in order to identify an unknown life distribution from a TTT plot by comparing it with possible alternative scaled TTT transforms (see Fig. 2).

Properties of the scaled TTT transform are given by Barlow & Campo (1975) (see also an expository paper, Bergman & Klefsjö 1984b, where further references are found). A central property, when F is a.c., is that

$$\frac{d\phi(u)}{du} = \frac{1}{r\{F^{-1}(u)\}}; \qquad 0\leq u\leq 1,$$

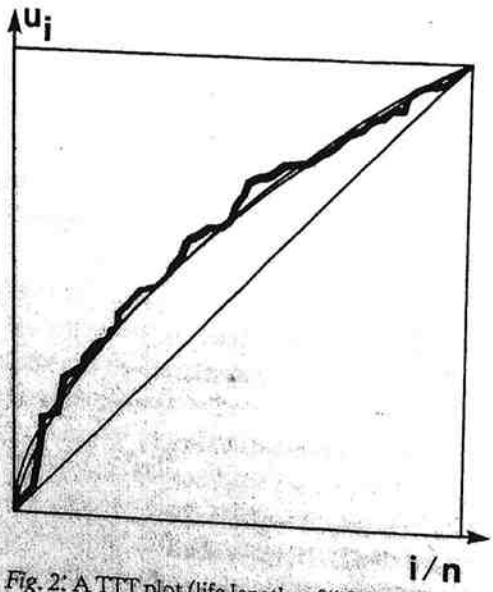

Fig. 2: A TTT plot (life lengths of "right rear brakes" (see Barlow & Campo, 1975)) together with Weibull and gamma TTT transforms (shape parameters 1.5 and 2.0, respectively).

which means that if e.g. F is IFR (DFR) then the scaled TTT transform is concave (convex). Using this and other properties we may characterize important non-parametric properties by using the scaled TTT transform (see e.g. Langberg *et al.*, 1980; Barlow, 1979; Bergman, 1979a; Klefsjö, 1982), and this may be used to design suitable test statistics based on the TTT plot (see e.g. Klefsjö, 1983). For instance, for testing exponentiality against IFR or, more generally, against NBUE alternatives (or their duals), the cumulative total time on test statistic, $\Sigma_{i=1}^{n} u_i$, is natural (c.f. Barlow *et al.*, 1972; Hollander & Proschan, 1975). In fact, this sum (multiplied by $1/n$) is essentially the area under the TTT plot and the corresponding area for the scaled TTT transform (-0.5) has been suggested by Barlow (1979) as a measure of "IFR-ness". The cumulative total time on test statistic has a weak optimality property (asymptotic maxi–min) for testing exponentiality against IFR or DFR alternatives (see Barlow & Doksum, 1972). It is very closely related to a trend test studied by Cox (1955); in his discussion of Cox (1955) paper, Bartholomew (1955) pointed out that this trend test was the oldest known statistical test having been developed by Laplace.

Observe that it is possible to use the TTT plotting technique also for censored data as long as the T_i's are calculated as the total time on test until the ith failure, and T_n the total time on test with respect to all observations. Under exponentiality u_i, $i=1,\ldots,n-1$, are still distributed as an ordered sample from a uniform distribution, and deviations from exponentiality may be detected as for complete samples (see Fig. 3).

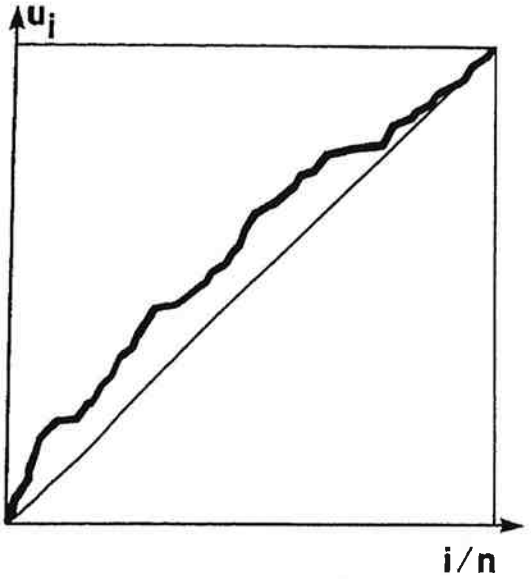

Fig. 3. TTT plot from an incomplete sample of times to first failure of 196 hydraulic cylinders (74 failed). The plot indicates that the failure rate is increasing, but we cannot identify the life distribution.

However, if we want to identify an unknown life distribution by use of the TTT plot from censored data, care has to be taken. We cannot, as in the case of complete data, use the convergence result mentioned earlier. In this case we have to use e.g. the Kaplan–Meier product-limit estimator of the life distribution (c.f. Kaplan & Meier, 1958) and we have to derive the corresponding empirical scaled TTT transform in order to be able to use the identification technique displayed in Fig. 2 (see e.g. Bergman & Klefsjö, 1984a).

3.3. Reliability data banks

Once we have chosen a model and checked it against data, we have to make proper inference from the data; Bayesian methods are appropriate in most design decision situations. In the

design and redesign process, we have to use earlier gathered data available in more or less general data banks both as a source for prior information and, in some cases, as the only basis for an evaluation.

As mentioned also by Lawless (1983), the collection (and analysis) of field data poses challenging and important problems, yet it has not been much discussed in the statistical literature.

As an example on methods for field data evaluations we shall present a method used at the Aerospace Division, Saab-Scania AB, for the evaluation of system and subsystem (component, unit) reliability (see also Bergman, 1979b).

Let the system under study be a series system with r components, each with an exponentially distributed time between failures; here we assume that each of the components is repaired to a new-like condition after failure and we neglect repair times. The system time between failures is exponential with parameter, failure rate, $\lambda_s=\Sigma_{i=1}^r \lambda_i$, where λ_i is the corresponding parameter of the ith component.

In early stages of the development process general handbook data (e.g. MIL-HDBK:217D) or similar data have to be used in order to assess system and component reliabilities. The result of this assessment, often called the "reliability prediction", probably gives more information on the relations between the failure rates of the components than on their absolute levels, since the failure rates in operational use may all be erroneous, e.g. because of a faulty assessment of the future environment, internal and external. But such faulty assessments tend to affect the failure rates of all components similarly, keeping relations among them rather stable. Observe that we have now been describing a type of dependence different from the shock model discussed in section 2. To describe system and component reliability characteristics we can use the vector $(\lambda_s, p_1, \ldots, p_r)$, where $p_i=\lambda_i/\lambda_s$. In the assessments based on handbook data and similar sources our uncertainty about λ_s is rather large, while the precision in $p_1, \ldots, p_r$ should be much better.

Now, assume that the system has been in operation for a while and that some failures have been obtained. Then we have gained a lot of information on λ_s, but we have not gained much information on $p_1, \ldots, p_r$.

Thus, we have two pieces of information complementing each other, and we should use both them to form our future assessments of system and component reliabilities. This may be done by using a prior distribution based on the handbook information and then updating this prior distribution using the obtained failure data. It is easily seen that a natural prior distribution in this case is for λ_s a gamma distribution with probability density function (p.d.f.) proportional to $\exp(-\lambda_s a)(\lambda_s a)^{p-1}$ and for $(p_1, \ldots, p_r)$ a Dirichlet distribution with p.d.f. proportional to $p_1^{\beta_1-1} \ldots p_r^{\beta_r-1}$. The parameters of the prior distributions have to be determined in accordance with past experiences about the precision of the used assessment technique and by use of so-called preposterior analysis; i.e. studies on the effect on the posterior distribution of different potential failure results. Typically, the means of λ_s, $p_1, \ldots, p_r$ (i.e. p/a and $\beta_i/\Sigma_{j=1}^r \beta_j$, $i=1, \ldots, r$) were chosen as the "predicted" values given by handbooks or other similar sources, while the shape parameter of the gamma distribution was of size 5 and $\Sigma_{j=1}^r \beta_j$ of size 2 or 3 times r, the number of components. Observe, that for this formalization of the prior information to be applicable, it is necessary that all the components (or subsystems) react on the environment or other affecting factors in the same way, i.e. that the system is homogeneous and the components are all (closely) related. In other cases it may be necessary to make an adjustment of the Dirichlet distribution so that it can model a more heterogeneous situation. A generalized Dirichlet distribution as the one suggested by Lochner (1975) could be used to model a situation where components within groups are closely related, while the relations between groups are less pronounced.

It should be observed that the situation modelled here is related to that studied by e.g. Mastran (1976) and Mastran & Singpurwalla (1978) for reliability assessments of coherent systems using both component and system test data. Thompson & Haynes (1980) have written a survey article on Bayesian interval estimation methods for system reliability.

3.4. Burn-in problems

As noted earlier, populations of electronic components often contain individuals having quality defects which considerably shorten component lifes. This is the argument for burn-in: a pre-usage operation of components performed in order to screen out the substandard components, i.e. those with shorter lives. Often the burn-in is performed in a severe environment, i.e. the time is accelerated. However, we shall not discuss time acceleration procedures here even if they are important from an applications point of view. If the burn-in period chosen is too short, some substandard quality components will still be put into operation and they will fail prematurely causing in some cases large consequence costs. On the other hand, if the burn-in period is too long, the burn-in costs will be high. An optimal burn-in period should be looked for.

Let τ be the random time to failure of a component and let F be its c.d.f. and μ its expectation. Also, assume that for each component exposed to burn-in we have a fixed cost a and a cost b per time unit during burn-in. Also, assume that c is the cost of a component and that K is the expected consequence cost of a failure in useful operation. Furthermore, let the optimality criterion be the expected long run cost per time unit; here we have used as an underlying assumption a renewal reward process in which the times between the failures (renewals) are the times of useful operation of the (burned-in) components (c.f. Ross, 1970). This cost is easily seen to be equal to

$$C_T=\frac{c+a+bE\{\min(T,\tau)\}+K\bar{F}(T)}{E\{\max(\tau-T,0)\}};$$

here T is the burn-in time. Observing that

$$E\{\min(T,\tau)\}=\int_0^T\{1-F(s)\}\,ds,$$

and

$$E\{\max(\tau-T,0)\}=\int_T^\infty\{1-F(s)\}\,ds,$$

$$C_T=\frac{K}{\mu}\,\frac{(c+a+b\mu\phi\{(F(T)\})/K+\bar{F}(T)}{1-\phi\{F(T)\}},$$

where as earlier $\phi(\cdot)$ is the scaled TTT transform of F. Let $u=F(T)$, and $C(u)=C_T$. If we can find an u_0 minimizing $C(u)$ we can also find an optimal burn-in period $T_0=F^{-1}(u_0)$. We have

$$C(u)=-b+\frac{K}{\mu}\,\frac{(c+a+b\mu)/K+1-u}{1-\phi(u)};$$

$$S(u)=\frac{1-\phi(u)}{d+1-u}$$

is maximized; here $d=(c+a+b\mu)/K$. But $S(u)$ may be illustrated graphically as the slope of the

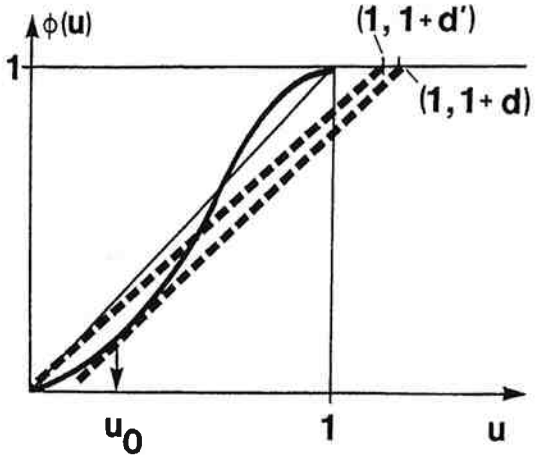

Fig. 4. An illustration of the graphical method for finding u_0, from which the optimal burn-in period may be determined.

line determined by the points $\{u, \phi(u)\}$ and $(1+d, 1)$. A graphical solution is easily obtained (see Fig. 4) and the optimal burn-in period can be determined.

Finally, we have to compare the cost of this solution with that without any burn-in at all. The latter cost corresponds to the slope of the line between the points $(0, 0)$ and $(1+d', 1)$, where $d'=(c+b\mu)/K$. If this slope is larger than $S(u_0)$ we shall not perform any burn-in. Observe, that this graphical solution makes it easy to perform sensitivity analyses with respect to, for example, different costs involved.

When the life distribution is unknown, but we have observed times to failures available, we can proceed in a similar way: the scaled TTT transform is replaced by the TTT plot and now it is easy to find an estimator of the optimal burn-in time by using the same technique suggested above (see also the illustration in Fig. 5).

Remark. In order to make bias corrections and also an approximate confidence interval for the optimal burn-in period we can use some resampling technique as e.g. boot-strapping (see Efron, 1982).

The above model, as well as some other burn-in models, are further studied in Bergman & Klefsjö (1985). The technique used is closely related to that of the determination of an

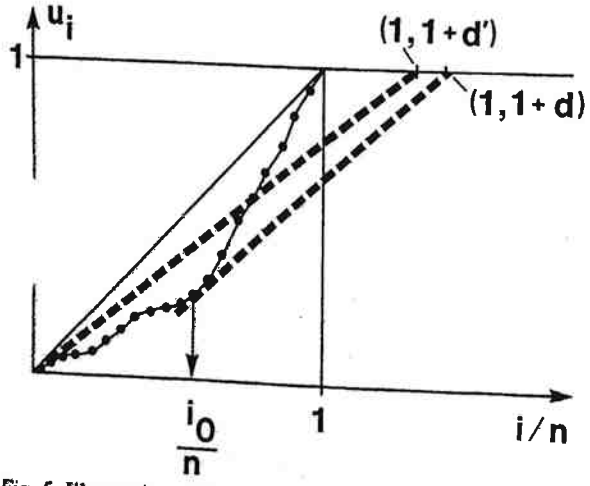

Fig. 5. Illustration of the estimation of the optimal burn-in time based on a TTT plot. Observe that we do estimate the optimal burn-in time to be larger than zero and in fact equal to $t_{(i_0)}$.

optimal age replacement interval (see e.g. Bergman, 1979a; Bergman & Klefsjö, 1982). The burn-in problem and the age replacement problem illustrate that the scaled TTT transform and the corresponding empirical transform, the TTT plot, display explicitly properties of interest for the solution of important reliability problems.

For a more general discussion of burn-in problems for repairable and non-repairable systems, see Jensen & Petersen (1983).

3.5. *Material strength models*

To find a quantitative measure of the reliability of a component we can use experiences from this or similar components in operation or from reliability life tests. However, sometimes a newly designed component is not similar to any other component in use. In such cases we may try to go further down into the component in order to infer reliability characteristics from strength properties of the material from which the component is built and from the stress properties of the environment. Areas in which a lot of work has been done are fatigue theory and fracture mechanics. However, the improvement of the corresponding probabilistic models and their use for reliability prediction and inspection strategies still poses challenging and important problems.

As an example (not in the above areas), on material strength modelling we shall study the modelling of composite material strength, an area in which considerable success has been made during the last years (see e.g. Smith, 1983), but where still much further research work has to be performed. A composite material consists of a number of lamina adhered to each other and each lamina consists of parallel fibres of e.g. carbon, kevlar or glass, within a matrix of e.g. epoxy. This material is, in relation to its weight, very strong and its failure is often connected with defects in e.g. fibres, their adherence to the matrix or in the adherence between the laminas. We shall restrict to the case of a unidirectional one-layer composite loaded in the direction of the fibres. A reasonable model for this type of material is the *chain-of-bundles* model, in which the material is looked upon as a chain of cross-sections, where each cross-section is assumed to have a short length δ determined by the properties of the fibre, the matrix and their adherence. Each cross-section is considered as a bundle of fibres. Failure of the material is obtained when the weakest cross-section has failed. Thus, the strength of a composite material is the strength of the weakest of a large number of bundles, each with a large number of fibres. Since large numbers of bundles and fibres are involved, asymptotic results should be applicable.

The strength of a single fibre is often assumed to be Weibull distributed. This seems quite reasonable since the failure of fibre is determined by its weakest part (largest defect) and thus its strength should be an extreme value distribution (see also Harlow *et al.*, 1983).

The strength of bundles of fibres with independent random strengths have been studied already by Daniels (1945), who proved that the asymptotic distribution of the bundle strength was normal under very weak assumptions on the fibre strength distribution. However, he assumed that the load corresponding to a broken fibre was shared equally by *all* fibres in the bundle and that the bundle failed when all of its fibres had failed. In a composite material the load of a broken fibre is redistributed on its closest neighbours, a local load-sharing rule (LLS) is in effect, and since other failure mechanisms are taking over after the failure of a number of adjacent fibres the bundle may be regarded as faulty as soon as e.g. k adjacent fibres are broken. In Bergman (1981a) and Smith (1983) it is shown that, if the load of a broken fibre is shared equally by its two closest neighbours, if the composite is regarded as failed if k adjacent fibres are broken, and if the fibre strength distribution has a Weibull-like left tail, then the bundle strength is (asymptotically) Weibull distributed with scale parameter determined by k

and the (tail properties of the) fibre strength distribution and the shape parameter determined by the corresponding shape parameter of the fibre strength distribution. This result is closely related to that of Loynes (1965) describing extreme value properties of uniformly mixing stationary processes.

Assuming independent random strengths of the bundles the composite material strength is seen to be (asymptotically) Weibull distributed and its scale and shape parameters may be found. Ultimately, this result should be useful not only for the prediction of the reliability of a composite structure but also for the control of size effects. For a general discussion of size effects and related important issues, not only with respect to composite materials, we refer to Harter (1977).

Still, a number of questions have to be answered concerning e.g. the adequacy of the proposed model (some experimental results support it), the possibility of modelling strengths in other directions than that of the fibres and of modelling the (random) strength of a material built by many lamina, in which failure may be a result of a combination of defects in the adherence between laminas and fibre defects.

Generally, care should be taken when reliability is calculated from strength–stress considerations. Often not all of the failure modes and potentially weak spots of the component can be investigated; are the "right ones" chosen? Numerical measures on basic material properties are based on results from special tests; are these results adequate for the material in hand, and have the same quality assurance measures been in effect? And so on. For a further discussion of related questions, see Österberg & Öfverbeck (1977).

3.6. One-shot devices

For a one-shot device its functioning is called upon just once, and to determine its reliability we have to describe its probability to function at demand. This probability may be dependent on many different factors such as e.g. age, temperature and so on.

Many one-shot devices are explosives and for this type of device high reliability and high safety may be contradictory; if we increase sensitivity in order to increase reliability, we also increase the risk of a not called for explosion. Below, we shall discuss the safety aspect of an electro-explosive device (EED) from some different points of view. The ideas given are also applicable with respect to reliability.

First, it seems plausible that for each individual EED a certain critical voltage exists such that if the EED is exposed to a voltage above this critical level then the EED explodes. From some extreme value arguments (see Bergman, 1982) it is reasonable to assume that the critical voltage for a random EED is Weibull distributed, at least if the population is homogeneous, i.e. if the production process is strictly under statistical control. But even if the Weibull model seems realistic, some objections may be posed. We have made an assumption that the production process is under strict statistical control, meaning that local properties, e.g. inhomogeneities, introduce much more variation among the EEDs than do the more low frequency disturbances such as batch to batch variation. This is a very restrictive assumption. If it is not quite fulfilled the model may still be adequate for a short sequence of produced EEDs but for a larger number of EEDs we may in fact have a mixture between different subsequences for each of which the Weibull model is adequate.

This means that for an EED randomly selected from a large production we may be forced to assume that its threshold value is distributed as a mixture of Weibull c.d.f.'s. Since a mixture of Weibull c.d.f.'s generally is not a Weibull c.d.f.'s, we have obtained a very clumsy model because the parameters in the Weibull c.d.f.'s as well as the mixing disribution are unknown. To circumvent this type of problem we recommend that the Weibull model is used but that its

validity is restricted only to the left tail, i.e. that part of the c.d.f. which is of interest for the safety assessment. This more general model should be adequate even if the production process does not quite fulfil the restrictive assumption of being under strict statistical control, since it is reasonable to approximate the tail of a mixture of Weibull c.d.f.'s by using a Weibull c.d.f., at least if the number of components in the mixture is small.

In order to give a complete safety assessment of an EED it is necessary not only to analyse the safety of an EED produced by a production process under (almost) strict statistical control, but we must also analyse the production process *per se*, especially the risk that the production process goes out of control in a way such that defective and unsafe EEDs are produced *and* also pass the quality control and eventually come into use. For this purpose we suggest the use of the fault tree analysis (FTA) technique as described earlier, here applied to the production process.

In order to perform the safety assessment of an EED produced in a production process under strict statistical control some statistical problems arise; from a limited number of firings of EEDs the parameters of the approximating Weibull c.d.f. have to be estimated. This is a difficult problem since only the first firing of each selected EED may be used; this means that the threshold level is not directly observable. If the EED exploded at the level used we only know that the threshold level of this EED was lower than the given level and we have the opposite information if it did not explode. We cannot reuse a fired, unexploded EED because the firing may have changed the characteristics of the EED even if it did not explode; this change of characteristics is often called "dudding". We have to consider each firing as a destructive test.

Two types of statistical problems have to be solved in order to perform the safety assessment:

1. Experimental design.
2. The estimation procedure.

Similar statistical problems arise in bioassay and in the determination of fatigue endurance limits; they are often called quantal response problems. Methods have been adopted for the normal and logistic distributions (cf. Finney, 1971; Berkson, 1953) and purely non-parametric methods have also been developed (see Ramsey, 1972). Here we shall indicate how Bayesian statistical analysis may be used in order to determine an experimental design and to estimate the parameters of the approximating c.d.f.

We assume that we have some, possibly vague, ideas about scale and shape parameters of the approximating c.d.f., and that these ideas may be expressed by using a bivariate density function $g'(a, b)$, $0<a<\infty$, $0<b<\infty$, the prior density function for the parameters α and β. After an observed firing at level x the prior density function is updated to give a posterior density:

$$g''(a, b) \propto \begin{cases} g'(a, b) \exp\{-(x/a)^b\}, & \text{no explosion;} \\ g'(a, b)[1-\exp\{-(x/a)^b\}], & \text{explosion;} \end{cases}$$

here the symbol $\propto$ means 'proportional to''. The level x is chosen so that, prior to the experiment, the expected increase in the Bayesian confidence limit of an extreme left tail percentile is maximized: Let $p=p(\alpha, a, b)$ be an $(1-\alpha)$ percentile of the c.d.f. with parameters a and b, and let $\bar{p}$ be a lower Bayesian confidence limit of p, i.e.

$$\int_{\{a,b|p(\alpha,a,b)>\bar{p}\}} g''(a, b)\, da\, db = P;$$

here P is the confidence degree. Obviously $\bar{p}$ depends on the prior density g', the level x and

the outcome of the experiment. It is reasonable to choose an x such that $\bar{p}=\bar{p}(x)$ has maximal expected value with respect to the prior distribution g'. However, levels which, according to the prior information, may be from the right tail of the approximating c.d.f. should be avoided. This procedure is repeated until the total number of firings devoted to the experiment are finished.

Sometimes we feel that we have much more information than we really want to use since we want to be able to defend the safety assessment against most critics. Then we may use what we shall call the "proposer–oppenent strategy". According to this strategy the estimation procedure uses a conservative prior density function (the prior of the opponent), but for the experimental design we use all available information formalized in a second prior density function (the prior of the proposer). Hence, for each firing we choose the level x such that the expected (with respect to the prior of the proposer) increase in the Bayesian confidence limit (utilizing the prior of the opponent) of an extreme left percentile is maximized.

4. Repairable units

In this section we shall study models for repairable units; here the concept "unit" may be interpreted as system, component, or part. A repair action may be initiated either from a preventive or a corrective maintenance task, and after the repair the unit may be renewed (*renewal models*), it may be in the state just before the failure initiating the repair action (*minimal repair*), or it may be in some intermediate state, i.e. not certainly either as new or as just before the failure (*incomplete repair*). We shall in this section also discuss some models applicable for systems, the failures of which are the consequences of inherent design faults existing within the system usually because of design errors, and which do not show up until certain environmental conditions are satisfied; the typical example is software, which is not physically changed because of environmental stresses, but still, it may fail because of an inherent fault and a consequential faulty response on certain input data combinations. When such a failure is experienced, a repair or rather *redesign action* is initiated, and hopefully the future reliability of the software is improved. Also other types of systems may experience the same type of reliability improvements.

4.1. *Renewal models*

When the repair action brings the unit to an as-new-like condition it is natural to model the failure process as a renewal process, where the renewal intervals are the times between successive failures. This model is also adequate if the state after a repair is the same irrespective of the number of earlier failures. In this case the first renewal interval may have a different distribution from the subsequent ones.

In this subsection we shall study two important applications of renewal theory: the calculation of availability characteristics and the determination of optimal replacement strategies.

Availability performance. Let $(\tau_1, \zeta_1), (\tau_2, \zeta_2) \ldots$ be a sequence of independent times to failure and repair times, i.e. ζ is the repair time of the ith failure; here repair time should be understood in a wide sense; also waiting times should be included, i.e. ζ_i is the complete downtime caused by the ith failure. Assume $E(\tau_i)<\infty$ and $E(\zeta_i)<\infty$, and let

$$U(t)=\begin{cases} 1, & \text{if } \sum_{i=1}^{\nu}(\tau_i+\zeta_i)\leq t<\sum_{i=1}^{\nu}(\tau_i+\zeta_i)+\tau_{\nu+1} \quad \text{for some } \nu\geq 0; \\ 0 & \text{otherwise;} \end{cases}$$

i.e. $U(t)=1$ if the unit is functioning at time t, and zero otherwise.

Then it is easily seen from general renewal theory that under weak (and natural) conditions

$$\lim_{t\to\infty} P\{U(t)=1\} = \frac{E(\tau)}{E(\tau)+E(\zeta)};$$

this is denoted limiting availability and usually the symbol A is used. Also the long run availability, i.e. the fraction of time in the long run during which the unit is available, is equal to A, i.e.

$$A = \lim_{t\to\infty} \int_0^t U(s)\, ds/t.$$

The importance of the availability measure A is self-evident. For an expensive system it is of great importance that it is operational. The gain of a productive system is directly proportional to its availability.

We may also define a momentary availability measure at time t, $A(t)=P\{U(t)=1\}$, but usually the convergence is rapid enough for the (limiting) availability to be sufficiently informative. For a more complete discussion and some further models close to the above one, see Barlow & Proschan (1981b). The above formula is applicable also when the given assumptions are not quite satisfied; see e.g. Franken & Streller (1978) for a case when independence is replaced by stationarity. But, in some applications this formula is used even if it should not be. Many systems are not intended to be used continuously but only during a part of the total time available. In many applications this seems to be forgotten. One model, which may be applicable, is a semi-Markow model, when e.g. three states are assumed: rest, operation and repair. In the semi-Markov approach we need to assume that times to failures are exponential and that also in other respects the future transition intervals are independent of the past. This is not the case if the times of operation are determined, e.g. from a queueing system, or if the total work of the unit is determined in advance. In these cases other availability models have to be built. In Gnedenko & Kovalenko (1968) some interesting availability models in queueing systems are studied, but otherwise this is an area in which more work has to be done.

During recent years buyers have realized the importance of good availability performance and sometimes they even force the vendors to guarantee the availability performance. To be "sure" that the guarantees are fulfilled statistical tests for use in availability demonstrations have to be derived. Some test plans are studied by Rise (1979) using properties of $\bar{U}(t)=\int_0^t U(s)\, ds/t$, in some special cases, as well as asymptotic properties. It was shown by Tacáks (1957) that under natural assumptions $\bar{U}(t)$ is asymptotically normal as $t\to\infty$ with mean A and standard deviation equal to $\sqrt{(\mu_\tau^2\sigma_\tau^2+\mu_\xi^2\sigma_\xi^2)/\{t(\mu_\tau+\mu_\xi)^3\}}$; μ_ξ and σ_ξ denotes expectation and standard deviation of a r.v. ξ. A somewhat more general result is given by Gut & Janson (1983).

Optimal replacement. The failure of a unit in operation may in many situations be costly and even dangerous. If we have reason to believe that older units are more prone to failure than younger ones it may be advantageous to replace a used unit by a new one at some age. In some situations we have information not only about the age of a unit under evaluation but also about its condition. If we have reason to believe that the proneness to failure depends on the condition of the unit, then it seems reasonable to use this in planning preventive maintenance.

We presume that a cost is associated with each replacement and that an additional cost is incurred at each failure in service. The problem is to find a good control strategy that balances the cost of replacements with the cost of failures and results in a minimum total long-run average cost per unit time or in a minimum expected total discounted cost.

In most cases we shall confine our study to the problem of minimizing the expected average

long run cost per time unit,

$$C_\delta = \frac{M(\delta)}{S(\delta)};$$

here $M(\delta)$ is the expected cost, and $S(\delta)$ is the expected time to replacement of an arbitrary unit, when a replacement rule δ is used. We have assumed that unit lives are independent. For a general result concerning the validity of the above formula, see Bergman (1980a). However, in some cases the same techniques may be used also to study expected total discounted costs, since in most cases this cost criterion may be written in a similar form:

$$\bar{C}_\delta = \frac{\bar{M}(\delta)}{\bar{S}(\delta)};$$

here $\bar{M}(\delta)$ is the expected discounted cost, $\bar{S}(\delta) = E\{1-\exp(-r\tau_\delta)\}$, r is the discount factor, and τ_δ is the stochastic time to replacement of an arbitrary unit if the replacement rule δ is used. Here we shall discuss some important aspects on replacement theory. For a general survey we refer to Pierskalla & Voelker (1975).

Assume that we know F, the c.d.f. of the time to failure of an arbitrary unit, and that we may choose a replacement rule based on age. For the replacement rule "replace at failure or at age T, whichever occurs first" the long run cost equals

$$C_T = \frac{c + K\,F(t)}{\int_0^T \{1-F(t)\}\,dt};$$

here c is the replacement cost and K is the additional cost incurred at failure (see e.g. Barlow & Proschan, 1965). It was noted in Bergman (1977) that the expression for C_T has some similarities with the scaled total time on test (TTT) transform discussed in section 3. In fact, the optimal age replacement interval T^* may be found by using the graphical procedure given in Fig. 6. Observe the similarity between this procedure and that of section 3.4, the determination of an optimal burn-in period.

This connection between the age replacement problem and the TTT transform made it possible to characterize those c.d.f.'s for which preventive age replacements might be economically feasible (see Bergman, 1979a). It also made it possible to find a very simple

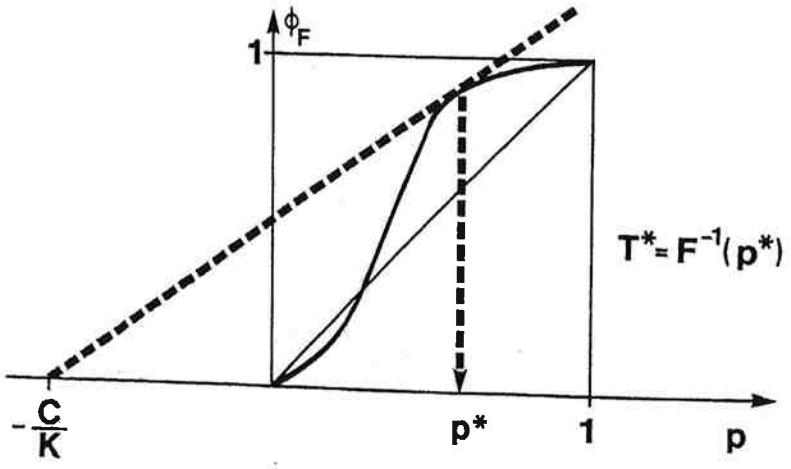

Fig. 6. A graphical procedure based on the scaled TTT transform for the determination of an optimal age replacement interval.

graphical procedure for the estimation of the optimal age replacement interval when the c.d.f. is unknown, but some observations on life lengths are available. In this case the scaled TTT transform in Fig. 6 is replaced by its empirical counterpart, the TTT plot (see Bergman, 1977). For some pleasant large sample properties of the obtained estimator see Bergman (1979a). In Bergman & Klefsjö (1983) the same technique has been used to study age replacement under a discounted cost model (see also Bergman & Klefsjö, 1984b).

We shall now study the problem of finding an optimal decision rule in problems where the cost criterion has the form $S(\delta)/M(\delta)$. Let A be the set in R^2 whose members may be written $x=\{S(\delta), M(\delta)\}$, $\delta\in D$, where D is the set of admissible decision rules, and denote as the "λ-problem" the problem of finding an admissible decision rule which minimizes the cost criterion $C_\delta^\lambda=M(\delta)-\lambda S(\delta)$. A result which has been used by e.g. Brender (see Barlow and Proschan, 1965) and Taylor (1975) in special cases is the following: if δ_λ minimizes C_δ^λ and $C_{\delta_\lambda}^\lambda=0$, then δ_λ minimizes C_δ. An illustration is given in Fig. 7, where $\delta_{\lambda'}$ minimizes $C_\delta^{\lambda'}$ and $C_{\delta_{\lambda'}}^{\lambda'}=0$. Thus $\delta_{\lambda'}$ also minimizes C_δ.

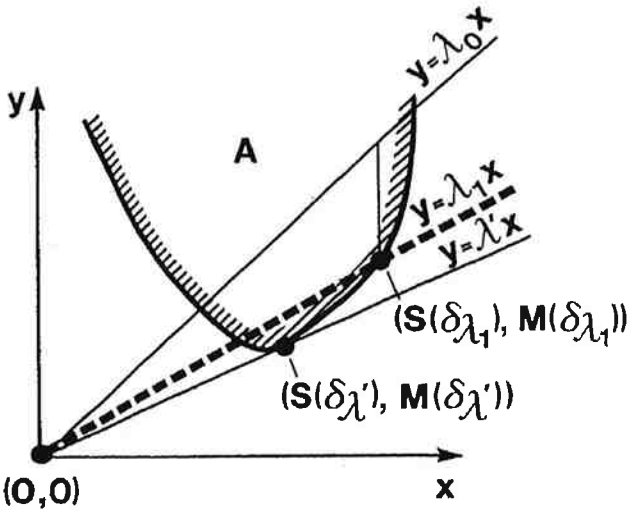

Fig. 7. Illustation of three λ problems, where the solution of the λ' problem is optimal with respect to long run cost per time unit.

Another result, which is very important, is the following: assume that to each λ-problem we can find an optimal rule δ_λ. Then the sequence λ_i, $i>0$, given by the algorithm $\lambda_i=M(\delta_{\lambda_{i-1}})/S(\delta_{\lambda_{i-1}})$, $i\geq 1$, where δ_{λ_0} is some arbitrary decision rule with $S(\delta_{\lambda_0})\neq 0$, converges to $\lambda'=\inf_{(\delta)} M(\delta)/S(\delta)$. If in Fig. 7 we continue to determine λ_2 by the given algorithm then it is easily seen that λ_2 is very close to the optimal λ'. The given algorithm is very important, since in most replacement problems it is rather easy to find solutions of λ-problems, and thus we can find approximate solutions. Indeed, it is shown in Bergman (1980) that if we do not confine ourselves to stationary replacement strategies, i.e. if we may choose different replacement rules for each new unit, then we can find an exactly optimal replacement strategy with respect to expected long run cost per time unit by using the suggested algorithm; for the first unit we choose δ_{λ_1}, for the second δ_{λ_2}, and so on. From a practical point of view we probably want to choose δ_{λ_0} such that λ_1 is close to λ', the optimal expected long run cost per time unit.

The same conclusion holds true even if we can find only ε-optimal solution to each λ-problem.

In a wide range of applications we know not only the age of a unit but we also have some measurements indicating its condition; these measurements may be wear characteristics, the

number of shocks incurred to the unit and so on. It seems natural to try to use this knowledge for the determination of an economic replacement strategy.

We presume the following model: let $X(t)$ be an observable stochastic characteristic of the unit at age t, and let $\nu(x)$ be the conditional failure rate of the unit given that $X(t)=x$, i.e.

$$P\{\zeta\in(t, t+dt)\,|\,X(t)=x, \zeta>t\}=\nu(x)\,dt+o(dt): \tag{4.1}$$

here ζ is the failure time of the unit and $o(h)/h\to 0$ when $h\to 0$. We assume that $\nu\{x(t)\}$ is increasing and right continuous with probability one. An admissable replacement rule may now be written "replace at failure or at the time τ whichever occurs first", here τ is a stopping time with respect to the stochastic process X, i.e. a random variable such that for any t the occurrence of the event $\{\tau\leqslant t\}$ is determined by the observation of $X(\cdot)$ up to time (age) t. Letting ζ be the age at failure we may write $C(\delta)$ as.

$$C_\tau=\frac{c+KP(\zeta\leqslant\tau)}{E\{\min(\zeta,\tau)\}};$$

here c is the cost of replacement and K is the additional consequence cost at failure. For the corresponding λ problem the cost criterion may be written

$$C_\tau^\lambda=c+KP(\zeta\leqslant\tau)-\lambda E\{\min(\zeta,\tau)\}.$$

After conditioning on $X(\cdot)$ we easily obtain

$$\begin{aligned}C_\tau^\lambda&=E\left[c+K\int_0^\tau \nu\{X(t)\}\,R(t)\,dt-\lambda\int_0^\tau R(t)\,dt\right]\\&=E\left[c+\int_0^\tau[K\nu\{X(t)\}-\lambda]\,R(t)\,dt\right];\end{aligned} \tag{4.2}$$

here $R(t)=\exp\{-\int_0^t \nu\{X(s)\}\,ds\}$. We have used the fact that if the process $X(\cdot)$ is known, then (4.1) describes an ordinary age-dependent failure rate function $\nu\{X(\cdot)\}$. From (4.2) it is easily seen that the replacement rule "replace at failure or as soon as $K\nu\{X(t)\}\geqslant\lambda$, whichever occurs first" is an optimal solution of the λ-problem. Thus the situation described earlier in this section applies, and it is easy to find the structure of the optimal replacement rule (take $\lambda=\inf_{(\tau)} C_\tau$). Also the algorithm described earlier is applicable.

The given model, essentially that given in Bergman (1978), may also be formulated in a more general point process context (see Aven, 1983a, b).

4.2. *Minimum repair models*

In this section we shall assume that the time τ to failure of a device under study has an a.c.c.d.f. F. In a statistical sense a minimum repair means that if a failure occurs at time t then, after the repair, the time to the next failure has distribution $\{F(t+s)-F(t)\}/\bar{F}(t)$ and failure rate function $r(t+s)$, $s\geqslant 0$. Thus, the number $N(t)$, of failures until time t, $t\geqslant 0$, is a non-homogeneous Poisson process with intensity function $\lambda(t)\equiv r(t)$; here we have assumed that repair times are negligible or that time is measured in operational time. However, some care has to be taken when this model is used; we have to distinguish between *statistical minimum repair*, for which the above interpretation is taken as a definition, and *physical minimum repair*, in which case the failed unit is restored to the exact physical condition as it was just before the failure. These two kinds of minimum repair are not necessarily equivalent! If the population, from which the device is taken, is not homogeneous, then each failure gives us some information on the subpopulation to which the device belongs. The distinction between the two minimum repair concepts does not seem to have been recognized earlier.

Minimum repair models have been studied first by Barlow & Hunter (1960). Recently, Aven (1983a, b) has studied minimum repair models in which more information than the age of the device is assumed available; the history of the device is assumed to determine a stochastic intensity function governing the point process of failures. This seems to be a fruitful approach for correct modelling of many natural situations in reliability applications as e.g. the modelling of physical minimum repair in a non-homogeneous population. In queueing theory, Brémaud (1981) has obtained nice results using the same type of approach.

Whenever non-homogeneous Poisson processes are encountered statistical methods are available (see e.g. Cox, 1955; Cox & Lewis, 1966). In some cases also obvious modifications of the TTT-plotting technique are applicable.

4.3. Imperfect repair models

In real-world situations the two earlier suggested repair models, complete and minimum repair, are two extremes which in many cases are good approximations but, usually, the result of a repair lies somewhere in between these two extremes. Some attempts to model this situation have been given.

In an interesting model suggested by Brown & Proschan (1983), it is assumed that the repair result is perfect with probability p and minimum with probability $q=1-p$. As shown by Brown & Proschan (1983) the distribution F_p of the time between two perfect repairs has failure rate function $pr(t)$, $t\geq 0$; here $r(t)$, $t\geq 0$, is the failure rate function of the c.d.f. F, the life distribution of the unit. From this fact important properties of F_p may be derived. Based on this model Fontenot & Proschan (1983) suggest some maintenance optimization models.

Another approach to imperfect repair is suggested by Nakagawa (1980) discussing several models in which the repaired unit never has effective age zero.

4.4. Software reliability

The widespread use of computers in systems has brought about the need for solutions of a lot of reliability problems not earlier encountered. The most challenging of these seems to be the modelling of software reliability. (Others are reliability improvements due to use of error correcting codes, the increased structural complexity and dynamic interactions.)

Software reliability models have been studied at least since the beginning of the seventies (see e.g. Jelinski & Moranda, 1972; Littlewood & Verall, 1973). Here we shall follow the approach of Littlewood (see e.g. Littlewood, 1980, and references cited there).

Based on failures experienced during the testing of a computer program we shall try to find some reliability characteristics of the program as e.g. the future failure times of the program; here time means execution time. The failures encountered during testing are the results of bugs (errors) inherent in the program, which manifest themselves once certain inputs are experienced, and after that, efforts are made to eliminate them. Let $t_1, \ldots, t_n$ be the experienced times between failures and let $T_{n+1}, T_{n+2}, \ldots$, be those not yet experienced. Based on $t_1, \ldots, t_n$ we want to predict $T_{n+1}, T_{n+2}, \ldots$, or some reliability measure derived from this sequence, as e.g. the probability of failure-free operation during a given period of time. If the debugging is successful then the random sequence $T_{n+1}, T_{n+2}, \ldots$, should be stochastically increasing. Its randomness originate from the fact that input data are chosen in a random fashion and that the bugs in the program are randomly distributed within the program; in fact, the program after the debugging following the nth failure divides the input space I in two parts: the set $I_F(n)$ of inputs for which a bug gives a failure and the corresponding complementary set. It seems fairly reasonable to assume that the time until an input data in

$I_F(n)$ is experienced is exponential with a parameter λ_n depending on $I_F(n)$. Since $I_F(n)$ is random, λ_n is also random. Littlewood & Verrall (1973, 1974) assume that $\lambda_1, \lambda_2, \ldots,$ is a sequence of independent r.v.'s with gamma distributions with probability densities

$$f_n\{\lambda; \alpha, \psi(n)\} = \frac{\psi(n)^\alpha \lambda^{\alpha-1} \exp\{-\psi(n)\lambda\}}{\Gamma(\alpha)};$$

here $\psi(n)$ is an increasing function of n in order $\lambda_1, \lambda_2, \ldots,$ to be stochastically decreasing and thus $T_1, T_2, \ldots,$ to be stochastically increasing. Since the distribution of T_n is a gamma mixture of exponential c.d.f.'s it is Pareto-distributed with probability density

$$f_{T_n}\{t; \alpha, \psi(n)\} = \frac{\alpha \psi(n)^\alpha}{\{t+\psi(n)\}^{\alpha+1}}.$$

In practice, $\psi(n)$ has to be chosen as a simple function of n, for example as $\psi(n)=\beta_1+\beta_2 n$, where β_1 and β_2 has to be estimated from the experienced times between failures $t_1, \ldots, t_n$ (see Littlewood, 1980).

As a natural extension of this and other models studied by Littlewood we suggest a stochastic point process approach, where the intensity function is determined by the experienced debugging and testing history of the program; formally, this history at time t corresponds to the sigma algebra of observable events until time t. This extension makes it possible to include also other relevant information concerning for example test effectiveness and the amount of reprogramming necessary to eliminate a bug; if a considerable reprogramming has to be made we can assume that the intensity is increasing instead of being decreasing.

In a way, testing software may be looked upon as having a function similar to hardware redundancies. Once we have found that the program gives a correct response for a certain set of inputs it will continue to do so (unless a hardware failure occurs). Attempts to implement redundant programs does not seem to be very effective; it is very expensive and failures may very well occur to both programs because of a faulty premise on which both programs have been built. It has been found that software failures experienced in operational use often originate from a faulty specification.

In order to improve the testing of a program it might be beneficial to let two (or more) independent groups test the program simultaneously. Probably this would increase testing effectiveness as well as the prediction of future reliability. A type of Bayesian capture–recapture model suggested by Jewell (1983) seems to be promising.

Remark. Hardware failures due to design errors may be looked upon in the same way. However, earlier used reliability growth models for hardware designs have been very crude; often non-homogeneous Poisson processes are suggested as models. The result of such models should be looked upon as being qualitative rather than quantitative. □

5. Acknowledgement

I want to thank Drs Bengt Klefsjö and Bent Natvig for many stimulating discussions on some of the topics treated in this paper. I also want to thank my former colleagues at Saab-Scania AB, Aircraft Division, for their support.

6. References

Aalen, O. O. (1975). Statistical inference for a family of counting processes. Ph.D. Thesis, Department of Statistics, University of California, Berkeley.

Aalen, O. O. (1978). Non-parametric inference for a family of counting processes. *Ann. Statist.* **6**, 701–726.

Aven, T. (1983a). Optimal replacement under a minimal repair strategy—a general failure model. *Adv. Appl. Prob.* **15**, 198–211.

Aven, T. (1983b). Contributions to failure time data analysis and optimal maintenance planning. PhD Diss., Institute of Mathematics, University of Oslo, Oslo.

Barlow, R. E. (1979). Geometry of the total time on test transform. *Naval Res. Logist. Q.* **26**, 393–402.

Barlow, R. E. (1982). Set theoretic signed domination for coherent structures. *ORC 82-1*. Operations Research Center, University of California, Berkeley.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. & Brunk, H. D. (1972). *Statistical inference under order restrictions*. John Wiley & Sons, New York.

Barlow, R. E. & Campo, R. (1975). Total time on test processes and applications to failure data analysis. In *Reliability and fault tree analysis* (ed. Barlow, Fussell & Singpurwalla). SIAM, Philadelphia.

Barlow, R. E. & Doksum, K. (1972). Isotonic tests for convex orderings. In *Proceedings of the 6th Berkeley symposium on mathematical statistics and probability*, Vol. 1, pp. 293–323. University of California Press.

Barlow, R. E. & Hunter, L. C. (1960). Optimum preventive maintenance policies. *Operat. Res.* **1**, 90–100.

Barlow, R. E. & Proschan, F. (1965). *Mathematical theory of reliability*. Wiley, New York.

Barlow, R. E. & Proschan, F. (1975). Importance of system components and fault tree events. *Stochastic Process. Appl.* **3**, 153–173.

Barlow, R. E. & Proschan, F. (1980). Inference for the exponential life distribution. Operations Research Center, University of California, Berkeley.

Barlow, R. E. & Proschan, F. (1981a). Life distribution models and incomplete data. Operations Research Center, University of California, Berkeley.

Barlow, R. E. & Proschan, F. (1981b). *Statistical theory of reliability and life testing*. To Begin With, Silver Spring, MD. (1st edn 1975)

Barlow, R. E. & Wu, A. S. (1978). Coherent systems with multistate components. *Math. Operat. Res*, **4**, 275–281.

Barholomew, D. J. (1955). Discussion of Cox (1955), *J. R. Statist. Soc. B* **17**, 162–163.

Bergman, B. (1977). Some graphical methods for maintenance planning. In *Proceedings, 1977 annual reliability and maintainability symposium*, Philadelphia.

Bergman, B. (1978). Optimal replacement under a general failure model. *Adv. Appl. Prob.* **10**, 431–451.

Bergman, B. (1979a). On age replacement and the total time on test concept. *Scand. J. Statist.* **6**, 161–168.

Bergman, B. (1979b). A Bayesian approach to system and subsystem reliability. In *Proceedings of the second national conference on reliability*, Birmingham, 1979, paper 4D12.

Bergman, B. (1980a). On the optimality of stationary replacement strategies. *J. Appl. Prob.* **17**, 178–186.

Bergman, B. (1980b). On some recent advances in replacement theory. In *Proceedings of the 6th advances in reliability technology seminar*, Vol 1, pp. 363–372.

Bergman, B. (1981a). On the probablity of failure in the chain-of-bundles model. *J. Comp. Mater.*, **15**, 92–98.

Bergman, B. (1982). On the safety assessment of electro–explosive devices. *Reliability Engineering* **3**, 193–202.

Bergman, B. & Klefsjö, B. (1982). A graphical method applicable to age-replacement problems. *IEEE Trans. Reliability* **R-31** (5), 478–481.

Bergman, B. & Klefsjö, B. (1983). TTT transforms and age replacements with discounted costs. *Naval Res. Logist. Q.* **30**, 631–639.

Bergman, B. & Klefsjö, B. (1984a). Total time on test transform. To appear in *Encyclopedia of statistics* (ed. Johnson & Kotz).

Bergman, B. & Klefsjö, B. (1984b). The total time on test concept and its use in reliability theory. *Operat. Res.***32**, 596–606.

Bergman, B. & Klefsjö, B. (1985). On burn-in problems and the TTT concept. To appear in *Quality and Reliability Engineering International*.

Berkson, J. A. (1953). A statistical precise and relatively simple method of estimating the bioassay with quantal response on the logistic function. *J. Am. Statist. Ass.* **48**, 565–599.

Birnbaum, Z. W. (1969). On the importance of different components in a multicomponent system. In *Multivariate analysis—II* (ed. P. R. Krishnaiah), pp. 581–592. Academic Press, New York.

Birnbaum, Z. W., Esary, J. D. & Saunders, S. (1961). Multicomponent systems and structures and their reliability. *Technometrics* **3**, 55–77.

Block, H. W. & Savits, T. H. (1981). Multivariate distributions in reliability theory and life testing. In *Statistical distributions in scientific work,* NATO Adv. Study Inst., Ser. C, Boston.

Brémaud, P. (1981). *Point processes and queues.* Springer-Verlag, Berlin.

Brown, M. & Proschan, F. (1983). Imperfect repair. To appear in *J. Appl. Prob.*

Calderola, L. (1980). Coherent systems with multistate components. *Nuclear Engineering and Design,* **58,** 127–139.

Cox, D. R. (1955). Some statistical methods connected with series events (with discussion). *J. R. Stat. Soc. B* **17,** 129–164.

Cox, D. R. (1972). Regression models and life tables (with discussion), *J. R. Stat. Soc. B* **34,** 187–202.

Cox, D. R. & Lewis, P. A. W. (1966). *The statistical analysis of series of events,* Methuen, London.

Daniels, H. E. (1945). The statistical theory of the strength of bundles of threads. *Proc. R. Soc. Lond. A* **183,** 404–435.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans.* SIAM, Philadelphia.

El-Neweihi, E., Proschan, F. & Sethuraman, J. (1978). Multistate coherent systems. *J. Appl. Prob.* **15,** 675–688.

Esary, J. D., Proschan, F. & Walkup, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38,** 1466–1474.

Finney, D. J. (1971). *Probit analysis,* 3rd edn. Cambridge University Press, London.

Fontenot, R. A. & Proschan, F. (1983). Some imperfect maintenance models. *Rep. No M667.* Florida State University, Tallahassee.

Franken, P. & Streller, A. (1978). A general method for calculation of stationary interval reliability of complex systems with repair. *EIK* **6,** 283–290.

Fussel, J. B. (1975). How to hand-calculate system reliability and safety characteristics. *IEEE Trans. Reliability* **24,** 169–174.

Fussel, J. B. & Veseley, W. E. (1972). A new methodology for obtaining cut sets for fault trees. *Am. Nuclear Soc. Trans.* **15**(1), 262–263.

Gnedenko, B. V. & Kovalenko, I. N. (1968). *Introduction to queueing theory,* IPST, Jerusalem.

Gut, A. & Janson, S. (1983). The limiting behaviour of certain stopped sums and some applications. *Scand. J. Statist.* **10,** 281–292.

Iaasl *et al.* (1981). *Fault tree handbook.* Office of Nuclear Regulatory Research, NUREG-0492, Washington, DC 20555.

Iahn, G. J. & Meeker, W. Q. (1982). Pitfalls and practical considerations in product life analysis, parts I and II. *J. Qual. Techn.* **14,** 144–152, 177–185.

Haines, A. & Singpurwalla, N. D. (1974). Some contributions to the stochastic characterization of wear. In *Reliability and biometry* (ed. F. Proschan and R. J. Serfling). SIAM, Philadelphia.

Harlow, D. G., Smith, R. L. & Taylor, H. M. (1983). Lower tail analysis of the distribution of the strength of load-sharing systems. *J. Appl. Prob.* **20,** 358–367.

Harter, H. L. (1977). Statistical problems in size effect on material strength. In *Applications of statistics* (ed. P. R. Krishnaiaha). North-Holland, Amsterdam.

Hollander, M. & Proschan, F. (1975). Tests for mean residual life. *Biometrika* **43,** 1136–1146.

Huseby, A. B. (1983). A generalized optimal factoring algorithm for computing exact system reliability, *Proceedings of the SRE symposium,* Arboga.

Huseby, A. B. (1984). A generalized domination theorem. To appear.

Jelinski, Z. & Moranda, P. B. (1972). Software reliability research. In *Statistical computer performance evaluation* (ed. W. Freiburger), pp. 465–484. Academic Press, New York.

Jensen, F. & Petersen, N. E. (1982). *Burn-in, an engineering approach to design and analysis of burn-in procedures,* Wiley, New York.

Jewell, W. S. (1983). Bayesian estimation of undetected errors. *Second Valencia international meeting on Bayesian statistics,* Valencia, Spain.

Joag-Dev, K. & Proschan, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11,** 286–295.

Kaplan, B. L. & Meier, P. (1958). Non-parametric estimation from incomplete observations. *J. Am. Statist. Ass.* **53,** 457–481.

Kaufmann, A., Grauchko, D. & Cruon, R. (1977). *Mathematical models for the study of the reliability of systems,* Academic Press, New York.

Keilson, J. (1979). *Markov chain models: exponentiality and rarity.* Springer, New York.

Klefsjö, B. (1982). On ageing properties and total time on test transforms. *Scand. J. Statist.* **9,** 37–41.

Klefsjö, B. (1983). Some test against ageing based on the total time on test transform. *Comm. Statist. A Theory and Methods* **12,** (8).

Langberg, N. A., Leone, R. V. & Proschan, F. (1980). Characterization of non-parametric classes of life distributions. *Ann. Prob.* **8,** 1163–1170.

Lawless, J. F. (1983). Statistical methods in reliability (with discussion). *Technometrics* **25,** 305–336.

Littlewood, B. (1980). Theories of software reliability. How good are they and how can they be improved? *IEEE Trans. Software Engineering* **SE-6** (5), 489–500.

Littlewood, B. & Verrall, J. L. (1973). A Bayesian reliability growth model for software reliability. In *Conference proceedings, 1973 IEEE symposium computer software reliability*, pp. 70–76, New York.

Littlewood, B. & Verrall, J. L. (1974). A Bayesian reliability model with stochastically monotone failure rate. *IEEE Trans. Reliability* **R-23,** 108–114.

Lochner, R. H. (1975). A generalized Dirichlet distribution in Bayesian life testing. *J. R. Statist. Soc.* **37,** 103–113.

Loynes, R. M. (1965). Extreme value in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36,** 993–999.

Mann, N. R., Schafer, R. E. & Singpurwalla, N. D. (1974). *Methods for statistical analysis of reliability and lifetime data.* Wiley, New York.

Martz, H. F. & Waller, R. A. (1982). *Bayesian reliability analysis.* Wiley, New York.

Marshall, A. W. & Olkin, I. (1967). A multivariate exponential distribution. *J. Am. Statist. Ass.* **62,** 30–44.

Mastran, D. V. (1976). Incorporating component and system test data into the same assessment; a Bayesian approach. *Operat. Res.* **24,** 491–499.

Mastran, D. V. & Singpurwalla, N. D. (1978). A Bayesian estimation of the reliability of coherent structures. *Operat. Res* **26,** 663–672.

Nakagawa, T. (1980). A summary of imperfect maintenance policies with minimal repair. *R.A.I.R.O.* **14,** 249–255.

Natvig, B. (1979). A suggestion of a new measure of importance of system components. *Stochastic Proc. Appl.* **9,** 319–330.

Natvig, B. (1980). Improved bounds for the availability and unavailability in a fixed time interval for systems of maintained, interdependent components. *Adv. Appl. Prob.* **2,** 200–221.

Natvig, B. (1982). On the reduction in remaining system lifetime due to the failure of a specific component. *J. Appl. Prob.* **19,** 642–652.

Natvig, B. (1982). Two suggestions of how to define a multistate coherent system. *Adv. Appl. Prob.* **14,** 434–455.

Natvig, B. (1984). Multistate coherent systems. *Encyclopaedia of statistical sciences*, Vol. 5. Wiley, New York.

Nelson, W. B. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics* **14,** 945–965.

Nelson, W. B. (1982). *Applied life data analysis.* Wiley, New York.

Pierskalla, W. P. & Voelker, J. A. (1975). A survey of maintenance models, the control and surveillance of deteriorating systems. *Naval Res. Logist Q.* **23,** 353–388.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics* **5,** 375–383.

Ramsey, F. L. A. (1972). Bayesian approach to bioassay. *Biometrics* **28,** 841–858.

Rise, J. (1979). Compliance test plans for reliability. In *Proceedings 1979 annual reliability and maintenance symposium*

Rosenthal, A. (1975). A computer scientist looks at reliability computations. In *Reliability and fault tree analysis*, (ed. R. E. Barlow, J. B. Fussell & N. D. Singpurwalla), pp. 133–152. (Discusses the computational complexity of fault tree analysis problems.)

Ross, S. M. (1970). *Applied probability models with optimization applications.* Holden-Day, San Francisco.

Räde, L. (1972). *Thinning of renewal point precesses.* Matematisk Statistik AB, Gothenburg.

Satyanarayana, A. & Chang, M. K. (1983). Network reliability and the factoring theorem. *Network* **13,** 107–120.

Satyanarayana, A. & Prabhakar, A. (1978). New topological formula and rapid algorithm for reliability analysis of complex networks. *IEEE Trans. Reliability* **R-27,** 82–100.

Smith, R. L. (1983). Limit theorems and approximations for the reliability of load-sharing systems. *Adv. Appl. Prob.* **15,** 304–330.

Takács, L. (1957). On certain sojourn time problems in the theory of stochastic processes. *Acta Math. Acad. Sci. Hungar.* **8,** 169–191.

Taylor, H. M. (1975). Optimal replacement under additive damage and other failure models. *Naval Res. Logist. Q.* **22,** 1–18.

Thompson, W. E. & Haynes, R. D. (1980). On the reliability availability and Bayes confidence intervals for multicomponent systems. *Naval Res. Logist.* **27**, 354–358.
Welsh, D. J. A. (1976). *Matroid theory*. Academic Press, New York.
Österberg, G. & Öfverbeck, P. (1977). *A study of outliers in statistical distributions of mechanical properties of structural steels*. Lund Institute of Technology.

Bo Bergman, Division of Quality Technology, Department of Mechanical Engineering, Institute of Technology, S-58183 Linköping, Sweden

DISCUSSION OF BO BERGMAN'S LECTURE

Elja Arjas (University of Oslo)

In the lecture Professor Bergman remarked on the *static* nature of today's system reliability calculations; the time variable t is usually given a fixed value and the evolution of the system reliability in time is ignored.

Except for time dependence, interdependence between the system's components is an important problem area which is often ignored. This can be the case, not only because arriving at natural mathematical formulations of interdependence is hard, but also because a realistic evaluation of the degree of dependence in any practical situation is likely to be extremely complicated.

I believe that time dependence and interdependence between the components should be considered within a single mathematical framework. Some support for such a view comes already from the fact that, in the *causality* relation, the cause necessarily precedes the effect in time. Although the convenient formulations of dependence in reliability theory are mostly weaker than actual causality (stochastic monotonicity seems to offer natural tools for considering such) the natural time ordering of the considered events remains a fundamental issue.

In the following I try to outline one possible approach to this part of reliability theory, which could be called *dynamic*. The mathematics is essentially that of "the martingale approach to point processes". This then forms one more link with the topics and techniques of the second invited paper on this meeting.

To take a simple example, consider an r component coherent system ϕ, where (i) there is no replacement of the components, and (ii) there are no multiple failures. Denoting by τ_i the life length of the ith component, the system life length τ_ϕ is an increasing function of the random vector $\tau=(\tau_i)_{1\le i\le r}$. Instead of considering the r dimensional probability distribution of τ, we look at the basic counting processes associated with the component failures: $N_i(t)=1-X_i(t)=1_{\{\tau_i\le t\}}$, $t\ge 0$. Thus $N_i(t)$ counts "one" at τ_i, jumping then from zero to one. Let $N_\phi(t)=1_{\{\tau_\phi\le t\}}$ be the corresponding counting process for the system failure.

To begin with, consider the notion of *hazard*. Suppose first that the behaviour of the system is monitored at *component level*. In other words, by time t the investigator knows, each component i, for which of the complementary events $\{\tau_i\le t\}$ or $\{\tau_i>t\}$ has occurred. If it is $\{\tau_i\le t\}$, he also knows the value of τ_i. The mathematical formulation of such knowledge is the generated history $\mathcal{F}_t=\sigma\{N_i(s);\ s\le t,\ 1\le i\le r\}$. Then, assuming absolute continuity, the stochastic $(\mathcal{F}_t)$ intensity $\{\lambda_i(t)\}_{t\ge 0}$ has the correct intuitive interpretation of an i component hazard rate: $\lambda_i(t)\,dt=P(\tau_i\in dt\,|\,\mathcal{F}_{t-})$. (The process $\{\lambda_i(t)\}_{t\ge 0}$ is defined implicitly by requiring that the difference $N_i(t)-\int_0^t\lambda_i(s)\,ds$ must be an $(\mathcal{F}_t)$ martingale.)

Note that, in agreement with the component level monitoring, $\lambda_i(t)=0$ for $t>\tau_i$. If the components are independent with respective distribution functions F_i and the failure rate functions are defined in the usual way by $r_i(t)=f_i(t)\{1-F_i(t)\}^{-1}$, we have $\lambda_i(t)=r_i(t)$ for $t\le\tau_i$.

The system hazard rate corresponding to component level monitoring can be defined analogously. It is easy to see that the system hazard rate is the sum of component hazard rates, where the summation, for each time t, is over the components critical at time t.

If the level of monitoring changes, the history must be changed accordingly. Thus, for example, if only the system failure can be recorded (but not the non-critical component failures), one must use the history $\mathcal{G}_t=\sigma\{N_\phi(s);\, s\le t\}$. It is then almost immediate that the corresponding system hazard rate, up to the system failure time τ_ϕ, is the same as the system failure rate function $r_\phi(t)=f_\phi(t)\{1-F_\phi(t)\}^{-1}$. ($F_\phi$ is the distribution function of τ_ϕ.) More interesting questions relating to the $(\mathcal{G}_t)$ filtration could concern state estimation, e.g. the estimation of the number and the positions of the failed components given that the system is still in the working condition.

The progressive growth of the knowledge $\mathcal{F}_t$ (or $\mathcal{G}_t$) in time t, as reflected above in the hazard rates, is but one aspect of the dynamic approach. More complete understanding about the dynamics of system reliability can be obtained by considering the sample path behaviour of the distribution valued process $\mu_t(\cdot)=P(\tau\in\cdot\,|\,\mathcal{F}_t)$, $t\ge 0$. Alternatively, one could consider, for suitably chosen functions $f:\mathbf{R}_+^r\to\mathbf{R}^1$, the time evolution of the conditional expected values $E\{f(\tau)\,|\,\mathcal{F}_t\}=\int f(\mathbf{x})\mu_t(d\mathbf{x})$, $t\ge 0$. For fixed f, the process $E\{f(\tau)\,|\,\mathcal{F}_t\}$ is obviously an $(\mathcal{F}_t)$ martingale. A natural choice would be $f(\tau)=\tau_\phi$, in which case this martingale describes how the estimated system life length evolves with t, always given the component level information up to t.

The following minor observation relates this martingale to the importance measure $I_N^{(i)}$ of Natvig, which was considered at the end of section 2. The construction of $I_N^{(i)}$ was based on the use of random variable Z_i whose definition becomes somewhat problematic and requires the notion of minimal repair if the components are not independent. However, denoting $M_\phi(t)=E(\tau_\phi\,|\,\mathcal{F}_t)$ and $\Delta M_\phi(\tau_i)=M_\phi(\tau_i)-M_\phi(\tau_i-)$, we see that $-\Delta M_\phi(\tau_i)$ has the interpretation: the reduction in the system life length estimate which is experienced upon the failure of the ith component. Thus $-E\{\Delta M_\phi(\tau_i)\}$ could be used in place of $E(Z_i)$, leading to a definition of $I_N^{(i)}$. A condition called WBF ("weakened by failures"), introduced in Arjas & Norros (1984), implies that $-\Delta M_\phi(\tau_i)\ge 0$, i.e. the failure of a component, is never beneficial to the system. (Here we have assumed that the exact time of occurrence of the ith component failure is "totally unpredictable", essentially meaning that the prediction does not place positive masses on individual time points. Otherwise a technical refinement to this definition is needed.)

The above example demonstrates how the basic stochastic processes arising from the dynamic approach can become instruments for considering dependence between components. Purely informally, one is led to consider ways in which $P(\tau\in\cdot\,|\,\mathcal{F}_t)(\omega)$ depends on ω (or rather, on the "$\tilde{t}$"-equivalence class in the sense of Jacobsen (1982)). It seems to me that such considerations could be mathematically attractive. More importantly, they could perhaps give some insight into dependence questions when complicated real systems are being considered.

References

Arjas, E. (1981a). A stochastic process approach to multivariate reliability systems: notions based on conditional stochastic order. *Math. Operat. Res.* **6**, 263–276.

Arjas, E. (1981b). The failure and hazard processes in multivariate reliability systems. *Math. Operat. Res.* **6**, 551–562.

Arjas, E. & Norros, I. (1984). Life lengths and association: a dynamic approach. *Math. Operat. Res.* **9**, 151–158.

Jacobsen, M. (1982). *Statistical analysis of counting processes.* Lecture Notes in Statistics. Springer-Verlag.

Norros, I. (1983). Systems weakened by failures. Submitted for publication.

Marvin Rausand (SINTEF)

Professor Bergman has given us an excellent introduction to the basic theory of reliability. Especially, I enjoyed the prologue to his paper. In my opinion, reliability must be considered as an interdisciplinary subject, somewhere in the intersection between mathematical statistics and a multitude of technological subjects. It is very important that the development in reliability theory is not self-contained and disconnected from the technological environment. The prologue of the paper, as well as Professor Bergman's background, opens the way for a sound approach.

In section 2 of his paper, Professor Bergman discusses systems reliability, mainly according to the same scheme as in Barlow & Proschan (1975). For a novice in this field, it may seem as structure functions based on reliability block diagrams are the only way of evaluating a system's reliability. I am sure that Professor Bergman did not want to leave this impression. In practical analyses, the structure of a system is usually established by means of a fault tree rather than by a reliability block diagram. Most of the computer codes available for determining the system's reliability are also written in the fault tree terminology.

There are also a number of other techniques, suitable for special situations, e.g. Markov/semi-Markov techniques and Monte Carlo simulation. Although normally classified as risk analysis techniques, the inductive techniques, like event tree analysis and cause–consequence diagrams, may be useful in many situations. I also want to mention the simple, but commonly used failure modes and effects analysis (FMEA).

Among the main problems encountered when we try to evaluate the reliability of a complex system are:

A number of different component failure modes, usually interdependent and competing to deteriorate the prime function of the components;
Non-constant failure rates;
Complicated and often undefined maintenance strategies, with a mixture of preventative and corrective maintenance;
Common cause/common mode failures;
Human reliability/operator behaviour;
Shortage of relevant reliability data.

Regarding common cause/common mode failures, Professor Bergman only reviews the method of associated variables. My experience is that this method leads to too wide bounds for the reliability, when used uncritically. There is now available a number of qualitative techniques suitable for identifying possible candidates for common mode failures. There is also available a number of other quantitative techniques to assess the effects of the dependencies.

When calculating a system's reliability, the input variables, like failure rates and repair times, will always be encumbered with some sort of uncertainty. These uncertainties, together with model short cuts, will usually lead to a relatively high degree of uncertainty in the estimate of the reliability of the system. Computer codes suitable for estimating this uncertainty have been developed. The first and most simple of these computer codes is the SAMPLE code, developed during the Reactor Safety Study, WASH 1400. This type of problem, which I consider to be very important, is not covered in Professor Bergman's paper.

Concerning measures of reliability importance, a topic thoroughly reviewed in the paper, I am sorry to say that till now I have not seen any significant use for these measures. Advocates for these measures claim that they may be used to come up with a priority list to identify failures in a system—a so-called fault-finding list. I doubt that any maintenance department

would trust such a list. One possible application of these measures might be in connection with the above-mentioned computer codes for estimating uncertainties. All these codes are based on Monte Carlo simulation and occupy a rather long computer time. The measures of importance could perhaps be used to make up an intelligent sequence for the simulation. Combined with some sort of a stopping rule, this could perhaps create more effective codes. (The idea is free.)

Section 3 of the paper is an excellent review of various techniques for evaluating the reliability of a component. At SINTEF, we have developed a set of computer codes to estimate reliability parameters for a component. Among other techniques, these codes cover the total time on test (TTT) transforms, the Kaplan–Meier estimates and the hazard plotting technique for multiple censored data. These codes have been used extensively on a wide range of data sets from off-shore equipment. My experience is that the hazard plotting technique is the best suitable technique to identify variations in the failure rate. The plot is also relatively easy to explain to engineers not familiar with statistical terms.

The Cox (1972) regression model is only briefly mentioned in the paper. I feel that this type of model will be very important in offshore applications in the near future, especially for subsea installations. Due to extremely high repair costs, the reliability of the equipment chosen for subsea installations will be of utmost importance. To be able to chose the best equipment, both regarding design and material selection, the oil companies have to perform some sort of accelerated life testing. The proportional hazards models, or Cox regression models, seem to be the best available models as a basis for such an accelerated life testing. They may, however, be a bit too simple to explain all the relevant combinations of failure mechanisms.

The last sections of Professor Bergman's paper give a very well-written survey of the theory, and should offer valuable information, both to the newcomer and the experienced reliability engineer.

I want to conclude my part of this discussion by congratulating Professor Bergman on an excellent introduction to a subject which I foresee will increase in importance during the next few years.

Bent Natvig (University of Oslo)

I am not sure if all of you are aware of the fact that Bo Bergman had been working in industry (Saab-Scania) for 15 years before taking up his present position as a professor. As for the future, I hope more professors in our subject will have such a background. For people like me, mostly working in the *theory* of reliability, it is very inspiring to have a discussion partner with a background in industry like Bo, very active also in basic research, who can tell whether our work is of practical relevance or not. I will congratulate him with a review given at this meeting, being more audience-oriented than is usual. This should not lead the audience to the false conclusion that the ideas presented in his talk are simple. A careful look at his written material and some of the references is a worthwhile investment to see the originality of a series of his points of view.

As a motivation for stochastic modelling in reliability, consider Fig. 1, presenting "a complete safety model" developed by the Norwegian safety offshore programme (Wulff, 1982). This program was running from 1978 to 1982 and costed 113 million NK. For all of you who find this model satisfactory, the rest of my contribution will be just a waste of time, and you do better in relaxing outdoors.

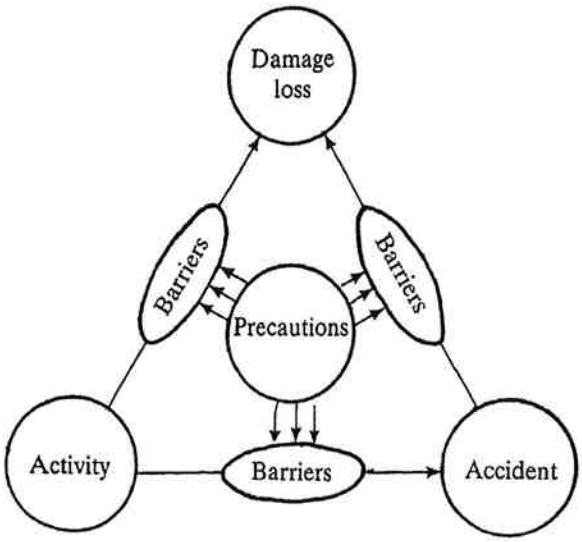

Fig. 1. A complete safety model.

Before going to the detailed comments I will just give the Society of Reliability Engineers (SRE)—Scandinavian Chapter—some publicity. This organization, having nearly 300 members from the Nordic countries, arranges a yearly symposium and has its own newsletter. It is a common platform for people doing risk analysis in industry and people working in reliability theory at the universities. The latter group is an influential minority. The two groups live together; of course not in complete harmony.

Turning to Bergman's talk, like Marvin Rausand I find the prologue excellent. However, I have one question on the LCC concept (and I know the answer). Does there exist a theory linked to this important and ambitious concept or is the industry mostly playing around with a terminology, a computer and some guessed costs?

I next turn to multistate reliability theory. Bergman suggests that the theory of ordered component/system states as introduced for instance in Natvig (1982, 1984) should be combined with the one of non-ordered states as presented in Caldarola (1980). This idea is obviously good. However, the latter paper is in my opinion not giving much more than a standard Boolean algebra treatment of the subject and not much theory is developed. I would nevertheless like to take up the challenge of combining the theories (hopefully on behalf of ıe of my students).

The multistate theory of ordered component/system states enables one for instance in a ıwer generation system to let the system state be the amount of power generated or in an oil ,ipeline network, as demonstrated in Natvig (1985a), the max flow one can get through the network. Let now $\{0, 1, \ldots, M\}$ be the set of states of the system, the $M+1$ states representing successive levels of performance ranging from the perfect functioning level M down to the complete failure level 0. Let, furthermore, $\mathbf{X}(s)$ be the vector of component states at time s and $\phi\{\mathbf{X}(s)\}$ the corresponding system state. The availability, $h^{(I)}(j, \phi)$, and the unavailability, $g^{(I)}(j, \phi)$, to level $j \in \{1, \ldots, M\}$ in the time interval I for a multistate monotone system with structure function ϕ are defined by:

$$h^{(I)}(j, \phi) = P[\phi\{\mathbf{X}(s)\} \geq j, \forall s \in I]$$

$$g^{(I)}(j, \phi) = P[\phi\{\mathbf{X}(s)\} < j, \forall s \in I].$$

Note that $h^{(I)}(j, \phi) + g^{(I)}(j, \phi) \leq 1$, with equality for the case $I = [t, t]$. In Funnemark & Natvig (1985) we arrive at bounds for $h^{(I)}(j, \phi)$ and $g^{(I)}(j, \phi)$, based on corresponding information on

the multistate components, thus generalizing Natvig (1980) covering the binary case $M=1$. The components are assumed to be maintained and interdependent. Such bounds are of great interest when trying to predict the performance process of the system, noting that exact expressions are obtainable just for trivial systems.

The best bounds are arrived at by using modular decompositions. Such bounds are also in fact the only ones of practical interest for large systems. The reason is that the bounds arrived at without using modular decompositions are based on all minimal path and cut vectors of the system. To arrive at these for a system of a large number of components seems impossible for computers of today. However, the number of components of each module, and the number of modules, may be chosen to be moderate, making it possible to arrive at the minimal path and cut vectors both of the organizing structure and of each module. The strategy is then to arrive at bounds for the availabilities and unavailabilities for the modules and inserting these into the bounds for the availabilities and unavailabilities for the organizing structure. This finally leads to improved bounds for the availabilities and unavailabilities for the system.

Even for the case where the marginal performance processes of the system's components are independent in I, the bounds are based on the assumption that each of these processes is associated in I. When these processes are Markovian, a convenient sufficient condition for this to hold, in terms of the transition intensities, is given in Hjort *et al.* (1985).

The papers above show how basic the concept of association is in dealing with dependence in reliability. Marvin Rausand has remarked that this may lead to poor bounds. This is true, but it just reflects the fact that precise statements in reliability are impossible when just having vague information on the dependencies coming into play, as is very usual in a risk analysis of a real-life system.

As a further indication of the fact that the multistate approach is taking over in reliability theory, Natvig & Streller (1984) give for systems with independently working and separately maintained components an explicit formula for the mean time which the system in steady state sojourns in states not below a fixed critical level. The theory for stationary and synchronous processes with an embedded point process is applied. The paper generalizes Franken & Streller (1980) treating the binary case. Finally, Streller (1980) generalizes the paper Takács (1957), mentioned by Bergman, to the multistate case. Let

$$\tilde{Z}(t)=\int_0^t \phi\{\mathbf{X}(s)\}\, ds$$

and let $\tilde{P}$ be the distribution of a suitable stationary process with an embedded point process. Then under specific conditions it is shown that

$$\tilde{P}\{\tilde{Z}(t)/t\geq j\}\approx\text{normal as } t\to\infty$$

We next turn to computer programs for fault tree analysis. Willie (1979) discusses computer-oriented methodology for deriving minimal path and cut sets associated with arbitrary fault trees leading to the program FTAP. The PAFT F77 program developed by Feo (1983) and the SAW program developed by Høgåsen (1984) are not based on minimal path and cut sets. All programs can deal with complementary events and all model dependence in fault trees by replicated basic events. It is worth mentioning that these programs have been developed at academic institutions.

Marvin Rausand seems to question whether measures of importance of system components are of any use. I still take the chance of making this a main topic. To me trying to assess the probability of failure of a large technological system is almost impossible. This is due to a poor and often irrelevant database, to little knowledge on human components and as stated earlier, vague information on the dependencies coming into play. This was clearly demonstrated in the

Reactor Safety Study (1975) on the safety of nuclear reactors in the USA. Hence the use of risk analysis to back political decisions on controversial safety issues is at least doubtful. If, however, a political decision is already made, risk analysis and reliability theory can contribute essentially to *improve* the safety of a system. This is for instance true for Norwegian offshore systems and Swedish nuclear reactors. As Bergman points out, measures of relative importance of each component to system reliability are basic tools when aiming at such improvements. Inspired by a recent correspondence with him I have tried to answer his question on which of the existing measures are the most suitable. For practical reasons this is published separately in Natvig (1985b). As a preliminary conclusion it seems that the Natvig (1979) measure is advantageous. Generalizations to the multistate state case are treated in Natvig (1985a). Finally, it should be admitted that the costs of improving the components are not entered into the measures reviewed by Bergman. Hence a continued research in this important field is needed.

Turning to statistical inference in reliability, I start by mentioning the papers (Viertl, 1980, 1983; Strelec & Viertl, 1982) on accelerated life testing. Let $F(t)$ be the life-time distribution of a device under usual stress and $F_a(t)$ in the accelerated mode. Then their basic assumption is that

$$F_a(t)=F\{a(t)\} \qquad t\geq 0,$$

where $a(t)\geq t$ for all $t\geq 0$ is called the acceleration function. Several papers by the same research group on this important topic have been produced and more are expected to come.

Secondly, the way of identifying an unknown life distribution by using the TTT plot in the case of censored data, suggested by Bergman, seems very good. As an alternative to the Kaplan–Meier estimator of the life distribution, Bengt Klefsjø has pointed out to me that an estimator introduced in Koul & Susarla (1980) could be used. Also the use of the TTT plot to find an estimator of the optimal burn-in time seems very useful.

As Bergman states, it is natural to take the Bayesian approach in reliability. One of more good reasons is the often poor database. In a way I hope more statisticians will start to look into Bayesian methods as unprejudiced as people in reliability. Specifically, I will challenge the people gathered at this conference working on life history data. In some applications at least the database does not cry after asymptotic theory based on the martingale central limit theorem. On the other hand there is a great need to develop more and better Bayesian methods. The two applications by Bergman on reliability databanks and on electro-explosive devices (bombs?) are very nice examples of what is needed. Another nice example is Barlow & Shor (1984). Here, in a Bayesian setting, a stopping rule, given data, is defined as informative relative to parameters of interest if it is random and statistically dependent on those parameters. Practical examples considered in detail illuminate the role of informative stopping rules and show how they may arise in practice.

As final examples of interesting topics in Bergman's paper, which he did not get time to cover in his talk, I strongly recommend his discussion of material strength models and of software reliability. Knowing Bo, I still think he has left something up his sleeves for a later talk, which we of course look forward to.

References

Barlow, R. E. & Shor, S. W. W. (1984). Informative stopping rules. *Techn. Rep. 1.* Operations Research Center, University of California, Berkeley.

Feo, T. A. (1983). PAFT F77. Program for the analysis of fault trees. *Techn. Rep. 14.* Operations Research Center, University of California, Berkeley.

Franken, P. & Streller, A. (1980). Reliability analysis of complex repairable systems by means of marked point processes. *J. Appl. Prob.* **17**, 154–167.

Funnemark, E. & Natvig, B. (1985). Bounds for the availabilities in a fixed time interval for multistate monotone systems. *Adv. Appl. Prob.* **17**. (To appear)

Hjort, N. L., Natvig, B. & Funnemark, E. (1985). The association in time of a Markov process with application to multistate reliability theory, *J. Appl. Prob.* **22**. (To appear)

Høgåsen, G. (1984). SAW, a program for the analysis of fault trees. *Techn. Rep.* Institute of Mathematics, University of Oslo.

Koul, H. L. & Susarla, V. (1980). Testing for new better than used in expectation with incomplete data. *J. Am. Statist. Ass.* **75**, 952–956.

Natvig, B. & Streller, A. (1984). The steady state behaviour of multistate monotone systems. *J. Appl. Prob.* **21**, 826–835.

Natvig, B. (1985a). Recent developments in multistate reliability theory. *Proceedings of the IUTAM symposium to the memory of Waloddi Weibull "Probabilistic methods in the mechanics of solids and structures"*, Stockholm, 19–21 June 1984. Springer-Verlag, Berlin.

Natvig, B. (1985b). New light on measures of importance of system components. *Scand. J. Statist.* **12**, 43–54.

Reactor Safety Study (1975). An assessment of accident risks in U.S. commercial nuclear power plants. *WASH-1400.* Nuclear Regulatory Commission, Washington, DC 20555.

Strelec, H. & Viertl, R. (1982). Estimation of acceleration functions in reliability theory under multicomponent stress. *Progr. Cybernetics Systems Res.* **10**, 455–459.

Streller, A. (1980). A generalization of cumulative processes. *Elektr. Informationsverarb. Kybern* **16**, 449–460.

Viertl, R. (1980). Acceleration functions in reliability theory. *Meth. Operat. Res.* **36**, 321–326.

Viertl, R. (1983). Nonlinear acceleration functions in life testing. *Meth. Operat. Res.* **47**, 115–122.

Willie, R. R. (1979). Computer-oriented methods for assessing reliability of complex systems. Ph.D. Thesis, Operations Research Center, University of California, Berkeley.

Wulff, E. (1982). Talk reviewed in *Teknisk Ukeblad, Oslo,* no. 51.

Kjell A. Doksum (University of California, Berkeley)

First, let me congratulate Bo Bergman for an excellent series of lectures. He has given an illuminating exposition of the important topics and problems in reliability theory. In particular, the section on graphical methods for model identification is interesting. Here I want to emphasize that it is important to have statistical methods that give information about how reliable these graphs are. The study of these methods is called (reliability)^2 theory. For instance, in their lectures on counting process models at this conference, Per Kragh Andersen and Ørnulf Borgan gave an illustration of a martingale. This graph illustrates "noice", but when looking at it, one could imagine that there is an increasing trend in the right tail. In this case it is known that this trend is not real but due to random fluctuations. However, with real data from an experiment, this would not be known, and one could become convinced that a random fluctuation is a real trend unless one has a method for determining the reliability of the graph.

One simple method consists of giving a simultaneous confidence band. For instance, in the case of the total time on test (TTT) plot, one such approximate band is obtained by simply adding $\pm c/\sqrt{n}$ to the TTT plot, where c is the $(1-\frac{1}{2}\alpha)$ quantile from the table of the distributon of the absolute value of a Brownian bridge. This band will have confidence coefficient approximately $(1-\alpha)$ (see Barlow & Campo, 1975).

Another plot, which is popular in statistics but which is not used often in reliability, is the quantile–quantile (Q–Q) plot. When checking whether a sample could have been generated by an exponential distribution, one would plot $[x_{(i)}, K^{-1}\{(i-\frac{1}{2})/n\}]$, $i=1,\ldots,n$ where $x_{(1)} \leq \ldots \leq x_{(n)}$ are the order statistics of the sample, and K^{-1} is the inverse of the exponential

distribution function. If the sample is from an exponential population, the above plot will "tend" to fall close to a straight line through the origin, while if the distributon has an increasing failure rate, then the shape of the plot will "tend" to be convex. In this case, a simultaneous confidence band is given by $K^{-1}\{F_n(t)\pm k_\alpha\}$, where F_n is the empirical distribution function and k_α is the $(1-\frac{1}{2}\alpha)$ quantile from the table of the one-sample Kolmogorov statistic with estimated scale parameter. For details on this and other methods see the expository paper by Doksum & Yandell (1984). Recently, Hjorth (1985) has obtained results that can be used for testing shapes other than the exponential.

References

Barlow, R. E. & Campo, R. (1975). Total time on test processes and applications to failure data analysis. *Reliability and fault tree analysis* (ed. R. E. Barlow, J. B. Fussell & N. D. Singpurwalla). SIAM, Philadelphia.

Doksum, K. A. & Yandell, B. (1984). Tests for exponentiality. *Handbook of statistics* (ed. P. R. Krishnaiah & P. K. Sen), Vol. 4. Elsevier Science Publishers.

Hjorth, N. L. (1985). Contribution to the discussion of Andersen and Borgan's "Counting process models for life history data: A review". *Scand. J. Statist.* **12**. (To appear)

.e **Schweder** (University of Oslo)

First I must declare myself an outsider to reliability theory. From that position I will make two comments. First, I shall briefly discuss the concept of cause and causality in the reliability context. Then I shall give a reminder: As other methodology, reliability theory may fall into misuse—a misuse which may lead to dangerous technological over-optimism.

Cause and causality are difficult, but indispensible notions. In the reliability context the outsider is mostly exposed to these notions when an accident has occurred and the cause of the accident is sought. As with the *Alexander Kielland* accident (a dwelling platform in the North Sea that turned over in March 1980, leading to the death of 123 men), it is the singular chain of events which led to the actual accident that is catching most of the attention: How did the fatal crack in one of the platform legs come about? It is of course important to investigate the particular run of events which led up to the accident, but it must be kept in mind that for large and complex systems any particular accident, also the one which occurred, will usually be singular in the sense that in advance it has a low probability of occurring. Since there is usually a tremendous lot of possible fatal chains of events, one must take care that the one which occurred is not overshadowing the rest. This is important, but hardly a new insight. What has struck me, however, is the conceptual difference between two separate types of causal investigations after an accident. Such an investigation may either be undertaken to learn about the system at hand such that improvements in safety can be made, or it may have the very different purpose of finding out whether neglect or criminal action is connected with the accident. For the latter purpose, the particulars of the events which occurred are essential in identifying the "criminal" and determining his "crime". For the more important purpose of improving safety, it is however the general aspect of the events which are of prime interest. Since the reliability theorist (and practitioner) is supposed to look forward, his causal analyses must be *prospective in time and general in scope*. The causal analysis of the criminal investigator should, however, be *retrospective in time and singular in scope*. This distinction is unfortunately not always observed, at least not in the public debate, and it seems to me that the

retrospective singular causal analysis usually is given too much weight. This may lead to preventive actions against a repetition of the accident which occurred, resulting in less than optimal gains in the overall reliability of the system.

My other comment is more political. In the public debate there are technological pessimists and there are optimists. The safety, and more generally the controllability of technological systems is of course a dividing issue. Partly because of the Rasmussen report (Reactor Safety Study, 1975), a pessimist like me gets the impression that reliability theory is sometimes improperly taken into use by the optimists in order to scientifically "prove" their case. "Reliability theory" is a nice positive name, and the concepts and methods of the theory are, and must be, "positive" and "constructive". This aspect of "positivity" will usually fit in well with the thinking of the optimists. When analysing a given system, the reliability theorist must however establish a model of the system which ideally should encompass all possible evil events, and which sufficiently realistically describes the structure of the system. Model building is not easy, particularly not for large systems like the national energy system, and grave errors are likely to be made. The dangerously optimistic Rasmussen report was not, however, produced by statisticians, and like other similar reports it has fortunately been rebutted by people from our profession (Lewis *et al.* 1978; Bergqvist *et al.*, 1978; see also Fishoff *et al.*, 1982; Natvig, 1982). It is of course not only reliability theory which may fall into misuse. It is always hard for the public to tell the liar from the statistician. And it is a special duty for the reliability people to help prevent that comfortable technological optimism is unduly nourished by bad use of their theory. And in the nuclear power debate, this responsibility seems to have been adequately met.

Public reports like the Rasmussen report are open to attack from the responsible scientist. More dangerous are, however, the secret reports, particularly those produced in the military headquarters. I will now turn to the question of nuclear war by accident. It is obvious to me that reliability theory is of great interest to the military men—and I hope they make good use of it in preventing unnecessary death and destruction. But from the behaviour of the men in the White House and in the Kremlin it seems to me that the military men and their supremes are guided by misconceptions as to the reliability and controllability of their war machines. I am in no position to claim that faulty use of reliability theory is a major cause of these misconceptions, but it seems to me a tremendously important but difficult responsibility for scientists and reliability theoreticians in particular, to make their voices heard in order to correct these misconceptions. This challenge has in my view unfortunately not been satisfactorily met.

Reliability theory—and even more, good and proper use of it—is important, and I wish to thank Professor Bergman for his good lectures on the subject.

References

Bergqvist, I. *et al.* (1978). *Hur säkert kan man veta något om olycksriskerna i komplicerade tekniska system?* Institut för kärnfysik, Lunds Tekniska Högskola.

Fishoff, B., Slovic, P. & Lichtenstein, S. (1982). Lay foibles and expert fables in judgement about risk. *American Statistician* **36**, 240–255.

Lewis, H. *et al.* (1978). *Risk assessment review group report for the US Nuclear Regulatory Commission.* Washington, DC 20555.

Natvig, B. (1982). Sannsynlighetsregning of samfunn. *Statistical Memoirs* Institute of Mathematics, University of Oslo.

Reactor Safety Study (1975). An assessment of accident risks in US commercial nuclear power plants. *WASH-1400.* Nuclear Regulatory Commission, Washington, DC 20555.

Reply by Bo Bergman

First, I want to thank the discussants for filling in some of the gaps in my exposition. Especially, I have in mind issues like common cause failures, human reliability, and uncertainty importance, as well as multistate availability results, computer programs for fault tree analysis (FTA), accelerated life testing, uncertainties in the TTT plot, and the dynamic point processes approach to reliability. Some of these issues will be discussed later.

Bent Natvig asked a question on life cycle costs—certaintly there is quite a lot more to be done in reliability theory to give this important concept a stable foundation, but already the concept is of great value in the industrial world. Usually the most important cost factors as e.g. operating costs are quite simple, while others as e.g. loss of production due to unavailability and costs of spare parts already has a valid theoretical basis (see e.g. Barlow & Proschan, 1981, Ch. 7). In general, however, overall optimization strategies are missing.

The uncertainty of system reliability calculations due to the uncertainty of the component reliabilities, was emphasized by Marvin Rausand. This is a classical type of problem closely connected to the type of problems treated in section 3.3 of my paper (see also Martz & Waller, 1982, for further references). Still, there is quite a lot to be done to find adequate methods for the evaluation of the importance of the component reliability uncertainties with respect to system reliability uncertainty; here the ideas given by Rausand are worth considering. A discussion of this type of question is given also by Bier (1983).

Still, I consider reliability importance measures a useful tool for finding weaknesses in a design and for guidance on where to allocate efforts on reliability improvements. But, together with Marvin Rausand, I question their use as a tool for fault finding. For the corrective maintenance it is important not only to restore a faulty system into a functioning state but also to restore all of its redundancies and for this purpose importance measures cannot be of any guidance. Natvig (1985) gives a thorough discussion on the choice of relative reliability importance measures. Here I will only give some short complementary comments postponing some further suggestions to a forthcoming paper. Notations and assumptions used by Natvig (1985) will be used below.

For the choice of a specific importance measure it is important to make clear in what sort of situation it is to be used. Undoubtedly, different situations call for different importance measures—for a satellite with a limited economic or scientific life length we should probably use another importance measure than for a system whose most important reliability characteristic is its expected life length.

With the discovery of (1.5) in Natvig (1985) many problems concerning importance measures may be solved. But I think a slight generalization of that formula is even more basic and important: let the ith component have distribution function G_i instead of F_i and let T_{i,G_i} be the corresponding system life length. As in Natvig (1985), T denotes system life length, when the components, assumed independent, have life lengths with distributions $(F_1, \ldots, F_r)$. Then it is easily seen from the derivation of (1.5) in Natvig (1985) that

$$E(T_{i,G_i})-E(T)=\int_0^\infty \{\bar{G}_i(t-\bar{F}_i(t)\}\, I_B^{(i)}(t)\, dt. \tag{1}$$

Let us look on some different choices of G_i. First, let G_i be the distribution of the time to the second failure in a minimal repair model with underlying distribution F_i. Then (1.5) is obtained. If instead we make a small reliability improvement replacing $F_i(t)$ by $G_i(t)=F_i(t-c)$ then a natural basis for an importance measure is given by

$$\lim_{c\to 0}\frac{1}{c}\int_0^\infty \{\bar{F}_i(t-c)-\bar{F}_i(t)\}\, I_B^{(i)}(t)\, dt,$$

which may be interpreted as a system expected life length improvement per unit component life length improvement. This is easily seen to be equivalent to the Barlow–Proschan reliability importance measure. However, this type of constant reliability improvement irrespective of actual life length does not seem to be easily obtained by the designer—from the designer's point of view a small scale change should be more natural. Thus G_i should be chosen as $F_i(t/c)$, $c>1$. Then, a natural basis for the importance measure is

$$\lim_{c\to 1}\frac{1}{c-1}\int_0^\infty \{\bar{F}_i(t/c)-\bar{F}_i(t)\}I_B^{(i)}(t)\,dt=\int_0^\infty tf(t)I_B^{(i)}(t)\,dt.$$

Also improvements corresponding to those suggested for $I_{N_2}^{(i)}$ are natural from a design point of view—note that using the above ideas it is possible to define a generalization of $I_{N_2}^{(i)}$ for any life distribution. Finally, given the above indicated importance measures, cost considerations may be easily incorporated—reliability improvement should be allocated to that component for which it is most cost efficient.

As noted by Kjell Doksum it is very important to sort out what might be just random fluctuations in a TTT plot. Using asymptotic results is then the first step. However, in many cases the asymptotic results are not quite adequate because of slow convergences—here resampling schemes as e.g. boot-strapping may be useful (see e.g. Efron, 1981).

As indicated already in section 4.2, I find the dynamic point processes approach to reliability very promising. The concepts introduced to us by Elja Arjas really do strengthen this conviction. I am sure that the "weakened by failure" and similar concepts will prove very useful in the future.

The points taken up by Tore Schweder are important and they always need to be thought upon. I agree that sometimes to much attention is given to singular cases. However, not too seldom the thorough investigation of the singular case gives us information on general, adverse circumstances with respect to the applied design process, operational usage or maintenance procedures, which makes it possible to improve safety. And that is, as I see it, the main objective of the safety analyst. Just as a scientist cannot prove a scientific hypothesis about nature—he can only corroborate the hypothesis, to use the language of Popper—a safety analyst cannot claim to have proven a system safe; he can only tell from his analyses that he has not found the system unsafe, i.e. that the *identified* sequences of events leading to dangerous events carries small probabilities.

References

Bier, V. M. (1983). A measure of uncertainty importance for components in fault trees. Ph.D. Diss., LiOS-TH-1277, MIT.

Natvig, B. (1985). New light on measures of importance of system components. *Scand. J. Statist.* **12**, 43–54.