ISBN 82-553-0735-4 No. 4
April 1991

USING EXPERT OPINIONS IN BAYESIAN ESTIMATION OF COMPONENT LIFETIMES IN A SHOCK MODEL - A GENERAL PREDICTIVE APPROACH

by

Jørund Gåsemyr and Bent Natvig

USING EXPERT OPINIONS IN BAYESIAN ESTIMATION OF COMPONENT LIFETIMES IN A SHOCK MODEL - A GENERAL PREDICTIVE APPROACH

Jørund Gåsemyr and Bent Natvig

Institute of Mathematics, University of Oslo

Abstract

In this paper combining the opinions of k experts about the lifetimes of n components of a binary system is considered. This problem has been treated in the single component case by Huseby (1986, 1988). Since the experts often share data, he argues that their assessments will typically be dependent and that this difficulty cannot be handled without making judgements concerning the underlying sources of information and to what extent these are available to each of the experts. In the former paper the information available to the experts is modeled as a set of observations $Y_1, \cdots, Y_m$. These observations are then reconstructed as far as possible from the information provided by the experts and used as a basis for the combined judgement. This is called the *retrospective* approach. In the latter paper, the uncertain quantity is modeled as a future observation, Y, from the same distribution as the Y_i's. This is called the *predictive* approach. For the case, $n > 1$, where each expert is giving opinions about more than one component, additional dependencies between the reliabilities of the components come into play. This is for instance true if two or more components are of similar type, are sharing a common environment or are exposed to common cause failures. For the case $n = 2$ both the retrospective and predictive approach are treated in Natvig (1990). In the present paper the predictive approach is considered for an arbitrary n and for an arbitrary overlapping of the observation sets from the different experts. The component lifetimes are assumed to have a multivariate exponential distribution of the Marshall-Olkin type. At the end of the paper it is shown how the joint distribution of the lifetimes of the n components can easily be updated in the case of getting real data.

1. Introduction

Consider, for a fixed point of time, t, a binary system like a nuclear power plant of n binary components. Let $(i = 1, \cdots, n)$:

$$X_i = \begin{cases} 1 & \text{if the ith component functions} \\ 0 & \text{otherwise,} \end{cases}$$

$$\underline{X} = (X_1, \cdots, X_n),$$

$$\phi(\underline{X}) = \begin{cases} 1 & \text{if the system functions} \\ 0 & \text{otherwise.} \end{cases}$$

Let furthermore:

$$E(X_i|p_i) = p_i = \text{ the reliability of the ith component,}$$
$$E(\phi(\underline{X})|h) = h = \text{ the reliability of the system.}$$

If we assume that $X_1, \cdots, X_n$ are independent given $\underline{p} = (p_1, \cdots, p_n)$, we write:

$$h = E(\phi(\underline{X})|\underline{p}) = h(\underline{p}).$$

Natvig and Eide (1987) assumed that the joint prior distribution of the reliabilities, before running any experiments on the component level, $\pi(\underline{p})$, can be written as:

$$\pi(\underline{p}) = \prod_{i=1}^{n} \pi_i(p_i), \tag{1.1}$$

where $\pi_i(p_i)$ is the prior marginal distribution of p_i, i.e. we assumed that the components have independent prior reliabilities. $\pi_i(p_i)$ describes our initial uncertainty in p_i, by for instance allocating most of the probability mass close to 1 indicating a very reliable component.

In this paper we assume that k experts will provide the information about the reliabilities of the components. Our work in this area generalizes papers by Huseby (1986, 1988) on the single component case. Since the experts often share data, he argues that their assessments will typically be dependent and that this difficulty cannot be handled without making judgements concerning the underlying sources of information and to what extent these are available to each of the experts. In the former paper the information available to the experts is modeled as a set of observations $Y_1, \cdots, Y_m$. These observations are then reconstructed as far as possible from the information provided by the experts and used as a basis for the combined judgement of a decision maker (DM). This is called the *retrospective* approach. In the latter paper, the uncertain quantity is modeled as a future observation, Y, from the same distribution as the Y_i's. This is called the *predictive* approach.

For the case, $n > 1$, where each expert is giving opinions about more than one component, additional dependencies between the reliabilities of the components come into play. This is for instance true if two or more components are of similar type, are sharing a common environment or are exposed to common cause failures. In the case of $X_1, \cdots, X_n$ independent given $\underline{p}$, and the lifetimes being exponentially distributed with unknown failure rates $\lambda_1, \cdots, \lambda_n$, this problem is considered by Lindley and Singpurwalla (1986). Then obviously:

$$p_i = \exp(-\lambda_i t), \quad i = 1, \cdots, n.$$

In the latter paper the jth expert, $j = 1, \cdots, k$, expresses his opinion about λ_i and hence of p_i in terms of a normal distribution for $\theta_i = ln\lambda_i, i = 1, \cdots, n$. He provides its mean m_{ji} and standard deviation s_{ji} but also ρ_{jir} being the personal correlation between θ_i and $\theta_r, j = 1, \cdots, k;\ i, r = 1, \cdots, n,\ i \neq r$. In addition the DM has to provide his personal correlations between the m_{ji}'s for fixed expert j and different components, for fixed component i and different experts and finally for both different experts and components. The great drawback of this approach is the difficulty of assessing these correlations directly without having an underlying model as in the papers by Huseby.

Lindley and Singpurwalla (1986) use an approximation technique suggested by Laplace, which has been pointed out to be quite good by Tierney and Kadane (1986) to arrive at the corresponding uncertainty in $h(\underline{p})$ for a parallel system of independent components. They claim that the results may easily be generalized to cover any coherent system of independent components. This is not true since representing a coherent system of independent components by a series-parallel structure introduces replicated components which of course are dependent. For details see the excellent textbook Barlow and Proschan (1975).

For the case $n = 2$ both the retrospective and the predictive approach are treated in Natvig (1990). In the present paper the predictive approach is considered for an arbitrary n and for an arbitrary overlapping of the data sets from the different experts. Here the uncertain quantities $(Z_1, Z_2, \cdots, Z_n)$ are the lifetimes of the components. These are assumed to have a multivariate exponential distribution of the Marshall-Olkin type. As an example of such a two component system consider a module of the lubricating system for the main bearings of a power station turbine presented in Christensen and Kongsö (1991). In this system an oil pump driven by an electromotor is delivering oil from a reservoir, the oil being cleaned by a filter on its way. The system is a series system of five components. These are the oil reservoir which is failing if there is no oil, the filter which is failing if it is blocked and the electromotor, the oil pump and a power cable all failing if broken. Our module of interest is the series system of the oil pump and the electromotor. In the paper above the lifetimes of these components are assumed independent and exponentially distributed, neglecting that they are sharing a common environment or may be exposed to common cause failures, which are basic assumptions in our research in this area. If in addition the oil pump and the electromotor are of a new design, there is no data to rely on in the beginning. Hence the best one can do as a start is to let experts help in specifying the joint distribution of the lifetimes of these components.

The paper is organized in the following way. In Section 2 we consider the prediction of the joint survival distribution of the components given the background information of the experts. This is in terms of m independent sets of survival times beyond specific time points for all components. Section 3 is devoted to the extraction of information from the experts. This is achieved by asking the experts about the simultaneous survival probabilities, beyond a certain time point u, for different sets of components. In Section 4 we discuss claims to test the consistency of the experts, whereas Section 5 gives the predictions after the extraction of information from them. Section 6 is dealing with an alternative approach, where the background information is in terms of data on the times to shocks instead of survival times of the components. Updated predictions based on real data are treated in Section 7. An advantage of the alternative approach in Section 6 is that the proper predictions can be arrived at in a fully Bayesian fashion as in Huseby (1988). Some final comments are eventually given in Section 8.

2. Prediction based on simultaneous survival times

The lifetimes of the n components are assumed by the DM to have a multivariate exponential distribution of the Marshall-Olkin type. For details on the properties of this distribution we refer to Barlow and Proschan (1975). It is best described by the following shock model.

The components may fail in two different ways. First of all the ith component may fail due to an individual shock. The time until such a failure occurs is denoted by $V_i, i = 1, \cdots, n$. Furthermore, the ith component may fail due to a common shock that necessarily destroys certain other components as well. It is supposed that there are p possible common shocks. Introduce $(r = 1, \cdots, p)$:

$$D_r = \text{the set of components destroyed by the rth common shock.}$$

Denote by V_{n+r} the time until the rth common shock occurs, $r = 1, \cdots, p$. The variables V_i are assumed to be independent given the hyperparameters $\theta_i, i = 1, \cdots, n+p$ and exponentially distributed. Here θ_i is the failure rate corresponding to V_i. Also introduce $(i = 1, \cdots, n)$:

$$E_i = \{r | i \epsilon D_r\} = \text{ the set of common shocks that destroys the ith component.}$$

Then clearly the lifetime of the ith component satisfies $(i = 1, \cdots, n)$:

$$Z_i = \min\{V_i, V_{n+r}\ r \epsilon E_i\}.$$

We now suppose that the background information of the experts, corresponding to their observation sets, is in terms of m independent sets of survival times beyond specific time points for all components; i.e.

$$\bigcap_{i=1}^{n} (Z_{il} > z_{il}), \quad l = 1, \cdots, m. \tag{2.1}$$

This turns out to be mathematically advantageous and is in opposition to Huseby (1988) where observed lifetimes represent the corresponding information. Now define:

$$\begin{aligned} v_{il} &= z_{il}, \quad i = 1, \cdots, n; l = 1, \cdots, m \\ v_{(n+r)l} &= \max\{z_{il} | i \epsilon D_r\}, r = 1, \cdots, p; l = 1, \cdots, m. \end{aligned} \tag{2.2}$$

The information (2.1) is clearly equivalent to:

$$\bigcap_{i=1}^{n+p} (V_{il} > v_{il}), \ l = 1, \cdots, m, \tag{2.3}$$

where the V_{il}'s are independent and exponentially distributed with failure rates $\theta_i, i = 1, \cdots, n+p; l = 1, \cdots, m$.

Note that if no shocks overlap the system consists of independent modules, each of which is either a single component or consists of a fixed number of components with life distributions described by the individual failure rates of the components and the common shock failure rate. Thus under these circumstances the model becomes much simpler and each submodule may be treated separately.

The DM assesses that the jth expert has access to information on the Z_{il}'s in (2.1) or equivalently on the V_{il}'s in (2.3) for l with indices in the set $A_j, j = 1, \cdots, k$. We have:

$$\bigcup_{j=1}^{k} A_j = \{1, \cdots, m\}.$$

Thus the number of elements in each set A_j reflects the opinion of the DM about the quality of the assessments of the jth expert, whereas his judgement of the dependencies between the experts is reflected by the degree of overlap between the sets.

As in Natvig (1990) we assume that the prior distributions of $\theta_i, i = 1, \cdots, n+p$ both for the DM and the jth expert are independent gamma distributions with shape and scale parameters respectively equal to (a_i, b_i) for the DM and (a_{ji}, b_{ji}) for the jth expert, $j = 1, \cdots, k; i = 1, \cdots, n+p$. It should be noted that in principle the jth expert should specify the (a_{ji}, b_{ji})'s prior to giving the background information corresponding to his observation sets. In practice these gamma distributions may be chosen as rather vague ones with the scale parameters close to zero, but with the shape parameters not too small. Introduce:

$$\begin{aligned} t_{ji} &= \sum_{l \epsilon A_j} v_{il} \quad , j = 1, \cdots, k; i = 1, \cdots, n+p \\ t_i &= \sum_{l=1}^{m} v_{il} \quad , i = 1, \cdots, n+p. \end{aligned} \tag{2.4}$$

Here t_{ji} is the total survival of components from the ith shock, corresponding to the information from the jth expert. t_i is similarly the total survival corresponding to the whole set of information. As in Natvig (1990) we now have by standard calculations involving Bayes theorem the following prediction of the joint survival distribution of the components:

$$\begin{aligned} & P[\bigcap_{i=1}^{n}(Z_i > z_i) | \bigcap_{l=1}^{m} \bigcap_{i=1}^{n}(Z_{il} > z_{il})] \\ &= P[\bigcap_{i=1}^{n+p}(V_i > v_i) | \bigcap_{l=1}^{m} \bigcap_{i=1}^{n+p}(V_{il} > v_{il})] \\ &= k \int_{\theta_1=0}^{\infty} \cdots \int_{\theta_{n+p}=0}^{\infty} P[\bigcap_{i=1}^{n+p}(V_i > v_i) | \theta_1, \cdots, \theta_{n+p}] \\ &\times P[\bigcap_{l=1}^{m} \bigcap_{i=1}^{n+p}(V_{il} > v_{il}) | \theta_1, \cdots, \theta_{n+p}] \prod_{i=1}^{n+p} \frac{b_i^{a_i} \theta_i^{a_i - 1}}{\Gamma(a_i)} \exp(-b_i \theta_i) d\theta_1, \cdots d\theta_{n+p} \\ &= \prod_{i=1}^{n+p} \left(\frac{b_i + t_i}{b_i + t_i + v_i}\right)^{a_i}. \end{aligned} \tag{2.5}$$

Here the v_i's are defined in analogy with (2.2) by suppressing the index l. The constant k is determined by noting that $v_i = 0, i = 1, \cdots, n+p$ give a joint survival probability of 1.

Similarly we get:

$$P[\bigcap_{i=1}^{n}(Z_i > z_i)| \bigcap_{l \epsilon A_j} \bigcap_{i=1}^{n}(Z_{il} > z_{il})] = \prod_{i=1}^{n+p}(\frac{b_{ji}+t_{ji}}{b_{ji}+t_{ji}+v_i})^{a_{ji}}. \tag{2.6}$$

3. Extraction of information from the experts

It is seen from (2.6) that the DM must ask the jth expert questions that give him information on the $n+p$ quantities t_{ji} and then fit the t_{ji}'s to the information obtained. The simplest way of doing this is to get a system of $n+p$ equations involving the t_{ji}'s and solve the system. This can be done in many different ways. However, the questions ought to reflect the dependence structure in the model in a natural way and should also be as easy as possible for the expert to answer. This can be achieved by asking the expert about the simultaneous survival probabilities, beyond a certain time point u, for all components that can be destroyed by a certain shock. By doing this both for individual and common shocks one ends up with the following $n+p$ probabilities:

$$\begin{aligned} p_{js} &= P[Z_s > u| \bigcap_{l \epsilon A_j} \bigcap_{i=1}^{n}(Z_{il} > z_{il})], s = 1, \cdots, n \\ p_{j(n+r)} &= P[\bigcap_{i \epsilon D_r}(Z_i > u)| \bigcap_{l \epsilon A_j} \bigcap_{i=1}^{n}(Z_{il} > z_{il})], r = 1, \cdots, p. \end{aligned} \tag{3.1}$$

For our two component module the jth expert can for instance be asked to specify the probabilities that the oil pump alone survives 5000 hours, that the electromotor alone survives 5000 hours and that finally both survive 5000 hours. The right hand sides in (3.1) follow from (2.6). Hence the DM has $n+p$ equations to determine the $n+p$ quantities t_{ji}.

This system of equations may be unsolvable in which case the more general approach to follow may be used. Now the DM defines the partitions $\{F_s, F_s^c\}, s = 1, \cdots, n+p$ of the set $\{1, \cdots, n\}$ and asks the jth expert to specify the following $n+p$ probabilities:

$$p_{js} = P[\bigcap_{l \epsilon F_s}(Z_l > u) \bigcap \bigcap_{l \epsilon F_s^c}(Z_l > 0)| \bigcap_{l \epsilon A_j} \bigcap_{i=1}^{n}(Z_{il} > z_{il})], s = 1, \cdots, n+p. \tag{3.2}$$

This reduces to (3.1) in the special case where F_s is running through the components corresponding to the individual and common shocks. Introduce ($s = 1, \cdots, n+p$):

$$\begin{aligned} H_s &= \{n+r | D_r \bigcap F_s \neq \emptyset\} \\ &= \text{the set of common shocks that} \\ &\quad \text{affects at least some component in } F_s. \end{aligned}$$

From (2.6) we then have the following system of equations ($s = 1, \cdots, n+p$):

$$p_{js} = \prod_{i \epsilon F_s \cup H_s} (\frac{b_{ji}+t_{ji}}{b_{ji}+t_{ji}+u})^{a_{ji}}. \tag{3.3}$$

This corresponds to solving the system:

$$\underline{M} \cdot \underline{\zeta} = \underline{\eta}. \tag{3.4}$$

Here $\underline{M}$ is an $(n+p) \times (n+p)$ matrix whose ith row has 1's for elements in the set $F_i \cup H_i$ and 0's otherwise, $\underline{\zeta}$ is a vector having ith coordinate $(i = 1, \cdots, n+p)$:

$$\zeta_i = a_{ji} \ln((b_{ji} + t_{ji})/(b_{ji} + t_{ji} + u)),$$

and $\underline{\eta}$ is a vector having ith coordinate $(i = 1, \cdots, n+p)$:

$$\eta_i = \ln p_{ji}. \tag{3.5}$$

Thus in order to solve (3.4) the DM must construct a set of partitions such that the corresponding matrix $\underline{M}$ is invertible. We then get $(i = 1, \cdots, n+p)$:

$$\begin{aligned}&(b_{ji} + t_{ji})/(b_{ji} + t_{ji} + u)) = [\exp((\underline{M}^{-1} \cdot \underline{\eta})_i)]^{a_{ji}^{-1}} \\ &= (\prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}})^{a_{ji}^{-1}}.\end{aligned} \tag{3.6}$$

Hence finally $(i = 1, \cdots, n+p)$:

$$t_{ji} = (\prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}})^{a_{ji}^{-1}} \cdot u/(1 - (\prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}})^{a_{ji}^{-1}}) - b_{ji}. \tag{3.7}$$

In the special case above the matrix $\underline{M}$ will be a symmetric block matrix

$$\begin{pmatrix} I^{n\times n} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$

where $I^{n\times n}$ is the $n \times n$ identity matrix.

Furthermore, as an example consider the case with only one common shock wiping out all components at the same time. By modular decomposition this case also covers the one with nonoverlapping shocks as well. Then:

$$\underline{M} = \begin{pmatrix} I^{n\times n} & \underline{1}^{n\times 1} \\ \underline{1}^{1\times n} & 1 \end{pmatrix}.$$

Hence:

$$\underline{M}^{-1} = \frac{1}{n-1} \begin{pmatrix} (n-1)I^{n\times n} - \underline{1}^{n\times n} & \underline{1}^{n\times 1} \\ \underline{1}^{1\times n} & -1 \end{pmatrix}.$$

From (3.7) we then arrive at:

$$t_{ji} = (p_{ji}(p_{j(n+1)}/\prod_{s=1}^{n} p_{js})^{\frac{1}{n-1}})^{a_{ji}^{-1}} \cdot u/[1 - (p_{ji}(p_{j(n+1)}/\prod_{s=1}^{n} p_{js})^{\frac{1}{n-1}})^{a_{ji}^{-1}}] - b_{ji}, i = 1, \cdots, n$$

$$t_{j(n+1)} = (\prod_{s=1}^{n} p_{js}/p_{j(n+1)})^{\frac{1}{n-1}a_{ji}^{-1}} \cdot u/[1-(\prod_{s=1}^{n} p_{js}/p_{j(n+1)})^{\frac{1}{n-1}a_{ji}^{-1}}] - b_{j(n+1)}. \tag{3.8}$$

(3.8) is a generalization of (2.10) in Natvig (1990).

4. Consistency of the experts

In order to increase the precision in the assessments of the t_{ji}'s the procedure in the previous section should be repeated for different values of u. Note that for each u there is no guarantee that we end up by a solution satisfying the obvious claim ($r = 1, \cdots, p$):

$$t_{j(n+r)} \geq \max\{t_{ji} | i \epsilon D_r\} \geq \min\{t_{ji} | i \epsilon D_r\} \geq 0. \tag{4.1}$$

Furthermore, even if all of the sets of equations give acceptable solutions these will in general be different. Ideally we should calculate a posterior distribution for $t_{ji}, i = 1, \cdots, n+p$ based on these solutions. However, as an approximation we will at the present stage of research suggest as Huseby (1988) that one should base the subsequent calculations for fixed $j = 1, \cdots, k$ and fixed $i = 1, \cdots, n+p$ on the *averages* of t_{ji}, for the sets of equations having acceptable solutions. If an expert provides a small number of acceptable solutions, he should perhaps better be dismissed.

We will now discuss when the solutions (3.7) satisfy the claim (4.1) of being a set of acceptable solutions, or in other words that the experts are consistent. This parallels the discussion in Natvig (1990).

First of all we have from (3.6) and (2.6) ($i = 1, \cdots, n+p$):

$$\begin{aligned} \prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}} &= [(b_{ji} + t_{ji})/(b_{ji} + t_{ji} + u)]^{a_{ji}} \\ &= P[V_i > u | \bigcap_{l \epsilon A_j} \bigcap_{i=1}^{n} (Z_{il} > z_{il})] = \pi_{ji}. \end{aligned} \tag{4.2}$$

Hence the left hand side is the survival probability beyond u from the ith shock updated after the assessments of the jth expert. From (3.7) it is seen that a necessary condition for t_{ji} to be nonnegative is that this probability is less than 1, i.e.

$$\prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}} = \exp((\underline{M}^{-1} \cdot \underline{\eta})_i) < 1, \ i = 1, \cdots, n+p.$$

This is equivalent to:

$$\underline{M}^{-1} \cdot \underline{\eta} < \underline{0}^{(n+p)\times 1}. \tag{4.3}$$

It is, however, easy to show that (4.3) is in fact true if and only if the jth expert's probability assessments are consistent with the multivariate exponential distribution of the Marshall-Olkin type. From (3.5) this is equivalent to:

$$\underline{\eta} = -\underline{M} \cdot \underline{\theta}^{(n+p)\times 1} \cdot u, \tag{4.4}$$

which ensures (4.3) to hold. If on the other hand (4.3) is true, choose

$$\underline{\theta}^{(n+p)\times 1} = -\underline{M}^{-1} \cdot \underline{\eta}/u,$$

which ensures (4.4) to hold.

From (3.7) a necessary and sufficient condition for the t_{ji}'s to be nonnegative is that $(i = 1, \cdots, n+p)$:

$$(\prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}})^{a_{ji}^{-1}} u/(1 - (\prod_{s=1}^{n+p} p_{js}^{M_{is}^{-1}})^{a_{ji}^{-1}}) \geq b_{ji}. \tag{4.5}$$

By (4.2) this is equivalent to $(i = 1, \cdots, n+p)$:

$$\pi_{ji} \geq [b_{ji}/(b_{ji} + u)]^{a_{ji}},$$

which means that the survival probability beyond u from the ith shock is larger when updated after the assessments of the jth expert than à priori. This is intuitively obvious since nonnegative t_{ji}'s mean more experience without failures. (4.5) is always true when b_{ji} is small and a_{ji} is not too small.

From (3.7) and (4.2) the leftmost inequality in (4.1) is equivalent to $(r = 1, \cdots, p; i \epsilon D_r)$:

$$[\pi_{j(n+r)}^{a_{j(n+r)}^{-1}} u/(1 - \pi_{j(n+r)}^{a_{j(n+r)}^{-1}}) - \pi_{ji}^{a_{ji}^{-1}} u/(1 - \pi_{ji}^{a_{ji}^{-1}})] + [b_{ji} - b_{j(n+r)}] \geq 0. \tag{4.6}$$

It follows from (4.2) and (3.6) that the first of the two summands is nonnegative if and only if $(r = 1, \cdots, p; i \epsilon D_r)$:

$$a_{j(n+r)}^{-1}(\underline{M}^{-1}\underline{\eta})_{(n+r)} \geq a_{ji}^{-1}(\underline{M}^{-1}\underline{\eta})_i. \tag{4.7}$$

If the jth expert's probability assessments are consistent with the multivariate exponential distribution of the Marshall-Olkin type, it follows from (4.4) that (4.7) is equivalent to $(r = 1, \cdots, p; i \epsilon D_r)$:

$$a_{j(n+r)}^{-1}\theta_{(n+r)} \leq a_{ji}^{-1}\theta_i. \tag{4.8}$$

Let us furthermore consider the case where the jth expert assesses that:

$$a_{ji} = a_{j1}, \quad i = 2, \cdots, n+p. \tag{4.9}$$

Then nonnegativity of the second summand in (4.6) means that the jth expert assesses that the prior mean of $\theta_{(n+r)}$ is not less than the prior mean of $\theta_i, i \epsilon D_r$. Anyway, this is always true for the vague gamma distribution. When (4.9) holds, (4.8) is true if the jth expert's probability assessments are consistent with $\theta_{(n+r)} \leq \theta_i, i \epsilon D_r$. Hence the conditions for nonnegativity are somewhat opposite for the two summands. This makes sense as is seen from the following discussion.

If the jth expert assesses that $\theta_{(n+r)} > \theta_i, i \epsilon D_r$, the first summand is negative. This must be compensated by a nonnegative second summand claiming that the prior mean of $\theta_{(n+r)}$ is not less than the prior mean of $\theta_i, i \epsilon D_r$, which is the same type of assessment. In particular b_{ji} may have to be significantly larger than 0, in opposition to a vague gamma distribution. If on the other hand the jth expert assesses that the prior mean of $\theta_{(n+r)}$ is less than the prior mean of $\theta_i, i\epsilon D_r$, the expert is forced to assess that $\theta_{(n+r)} \leq \theta_i, i\epsilon D_r$. Finally, if the jth expert assesses that the prior mean of $\theta_{(n+r)}$ is not less than the prior mean of $\theta_i, i\epsilon D_r$, the expert may change his opinion without being inconsistent.

5. Prediction after extraction of information from the experts

What remains is for the DM to calculate the predictive joint survival distribution of the components given by (2.5) based on the extraction of the information $t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n+p$ from the experts. As is seen from (2.5) this can be done if the quantities t_i can be calculated by means of the t_{ji}'s. In the case where the experts are independent, i.e. where the sets A_j are disjoint, this is straightforward since from (2.4):

$$t_i = \sum_{j=1}^{k} t_{ji} \quad , i = 1, \cdots, n+p.$$

In general there exists a disjoint partition $B_g, g = 1, \cdots, q$ of the set $\{1, \cdots, m\}$ and subsets $C_j, j = 1, \cdots, k$ of the set $\{1, \cdots, q\}$ such that we have the representation:

$$A_j = \bigcup_{g\epsilon C_j} B_g \quad , j = 1, \cdots, k. \tag{5.1}$$

Let us define:

$$s_{gi} = \sum_{l\epsilon B_g} v_{il} \quad , g = 1, \cdots, q; i = 1, \cdots, n+p. \tag{5.2}$$

Then from (2.4) we have:

$$\begin{aligned} t_{ji} &= \sum_{g\epsilon C_j} s_{gi} \quad , j = 1, \cdots, k; i = 1, \cdots, n+p \\ t_i &= \sum_{g=1}^{q} s_{gi}, i = 1, \cdots, n+p. \end{aligned} \tag{5.3}$$

Hence the problem is solved if the s_{gi}'s were known. Note in particular that for a fixed partition $B_g, g = 1, \cdots, k$ of the data set it is irrelevant how many experts that share the information corresponding to each B_g as long as it is known by at least one expert. This is as expected since the DM doesn't assume errors in the observations by the experts. Of course, from the point of view of modelling dependencies the DM's predictive probability will decrease if the degree of overlapping between the A_j's increases, given a fixed set of t_{ji}'s.

The DM's final assessment of the predictive probability will be a result of his assessments of the s_{gi}'s. We start by considering these assessments in the special case considered by Huseby (1988) and Natvig (1990) where all experts share some common information, but otherwise have separate sources of information. Hence we have:

$$A_j = B_j \bigcup B_{k+1} \quad , j = 1, \cdots, k, \tag{5.4}$$

where B_{k+1} is the set corresponding to the common information. From (5.3) we then get:

$$\begin{aligned} t_i &= \sum_{g=1}^{k+1} s_{gi} = \sum_{g=1}^{k} (s_{gi} + s_{(k+1)i}) - (k-1)s_{(k+1)i} \\ &= \sum_{j=1}^{k} t_{ji} - (k-1)s_{(k+1)i}, \end{aligned} \tag{5.5}$$

where the unknown quantities $s_{(k+1)i}, i = 1, \cdots, n+p$ must satisfy:

$$0 \leq s_{(k+1)i} \leq \min\{t_{ji} | j = 1, \cdots, k\}. \tag{5.6}$$

This corresponds to (3.9) of Huseby (1988) and to (2.13) of Natvig (1990).

There are in principal several different procedures to assess the $s_{(k+1)i}$'s. First of all the DM may ask the experts of additional information. In Natvig (1990) this is done by simply letting $s_{(k+1)i}$ be the minimum survival time before the ith shock that all experts agree on. Secondly, the DM may decide that the information available only justifies calculating the two extreme versions of (2.5) corresponding to the two extreme values of (5.6). Suggesting the upper point as in Natvig (1990) gives the smaller predictive probability whereas $s_{(k+1)i} = 0$ gives the most optimistic assessment. Finally, the DM may use the information available to estimate the $s_{(k+1)i}$'s. Suggesting the upper point of (5.6) can be regarded as an estimate assuming maximal dependency within the model (5.4). Note, however, that the information about the number of elements in each of the sets B_g has not been used. This can be considered as known according to the general model and expresses the DM's judgement of the quality of the experts assessments and the dependencies between them. In the rest of this section we will consider estimation of the $s_{(k+1)i}$'s based on this information.

Remember that v_{il} introduced in (2.2) is the survival time from the ith shock coming from the lth observation set, $i = 1, \cdots, n+p; l = 1, \cdots, m$. We stress that the DM must assess that it is the same censoring mechanism behind all observation sets implying that the DM for fixed i must give equal weights to $v_{il}, l = 1, \cdots, m$.

Introduce:

$$\begin{aligned} n_g &= \text{ number of elements in } B_g, g = 1, \cdots, k \\ k_j &= \text{ number of elements in } A_j, j = 1, \cdots, k. \end{aligned}$$

We then suggest the following estimator for $s_{(k+1)i}$:

$$\hat{s}_{(k+1)i} = n_{(k+1)} \sum_{j=1}^{k} t_{ji} / \sum_{j=1}^{k} k_j. \tag{5.7}$$

In the general case (5.1) we have for fixed i to estimate s_{gi} for $q-k$ values of g. Then the remaining s_{gi}'s are determined from the first equation in (5.3). In the special case above the choice of estimating $s_{(k+1)i}$ was almost canonical, and there may exist an equally natural choice in other cases as well. Generally, it seems natural to estimate those s_{gi}'s for which the amount of available information is as large as possible. Define:

$$G_g = \{j | B_g \subset A_j\}. \tag{5.8}$$

Now pick the $q-k$ values of g for which the number of elements in G_g is the largest or alternatively for which:

$$\sum_{j\epsilon G_g} k_j$$

is the largest. For these values of g the following estimator is suggested:

$$\hat{s}_{gi} = n_g \sum_{j\epsilon G_g} t_{ji} / \sum_{j\epsilon G_g} k_j. \tag{5.9}$$

There are some obvious restrictions that must be obeyed in this estimation procedure. First of all if $B_g = A_j$ for some j, the index g must of course be excluded since then $s_{gi} = t_{ji}$ is already known. Similarly g must be excluded if B_g represents the only unestimated part of some A_j. Secondly, we must obviously have:

$$\sum_{g\epsilon C'_j} \hat{s}_{gi} < t_{ji},$$

where C'_j is the set of indices in C_j that is selected for estimation.

6. An alternative approach

In Section 2 the background information of the experts, corresponding to their imaginary observation sets, is given in terms of m independent sets of survival times beyond specific time points for all components, see (2.1) or equivalently (2.3). As has been seen this approach was mathematically tractable. In the case where k_j, the number of elements in $A_j, j = 1, \cdots, k$ are known, an alternative approach is possible. We then replace the background information (2.3) by:

$$\bigcap_{i=1}^{n+p} (V_{il} = v_{il}),\ l = 1, \cdots, m. \tag{6.1}$$

Hence we now have data on the times to shocks instead of survival times of the components.

It is of course generally impossible to observe the times to all shocks when observing a real system. However, the imaginary data are just an abstraction used to model the background information of the experts and do not have to be easily interpreted as real data. In fact we may in the following easily replace (6.1) by:

$$\bigcap_{i=1}^{n+p} \bigcap_{l=1}^{m_i} (V_{il} = v_{il}), \tag{6.2}$$

corresponding to a lack of information on some shocks in some data sets, to increase the flexibility of the model. It should be noted that also the imaginary data of (2.1) are unrealistic when failures of components in a real system are observed.

By defining the v_i's from th z_i's in analogy with (2.2) by suppressing the index l, and the t_{ji}'s and t_i's from the v_{il}'s as in (2.4), we get similarly to (2.5) and (2.6):

$$P[\bigcap_{i=1}^{n}(Z_i > z_i) | \bigcap_{l=1}^{m}\bigcap_{i=1}^{n+p}(V_{il} = v_{il})] = \prod_{i=1}^{n+p}(\frac{b_i + t_i}{b_i + t_i + v_i})^{a_i+m} \tag{6.3}$$

$$P[\bigcap_{i=1}^{n}(Z_i > z_i) | \bigcap_{l\epsilon A_j}\bigcap_{i=1}^{n+p}(V_{il} = v_{il})] = \prod_{i=1}^{n+p}(\frac{b_{ji} + t_{ji}}{b_{ji} + t_{ji} + v_i})^{a_{ji}+k_j}. \tag{6.4}$$

$a_i + m$ in (6.3) and $a_{ji} + k_j$ in (6.4) act as a_i and a_{ji} in respectively (2.5) and (2.6). Hence the choice of the shape parameters is less crucial in this approach. The procedure for extraction of the information from the experts are exactly as in Section 3. So the jth expert has to specify the $n + p$ probabilities ($s = 1, \cdots, n + p$):

$$p_{js} = P[\bigcap_{l\epsilon F_s}(Z_l > u)\bigcap\bigcap_{l\epsilon F_s^c}(Z_l > 0) | \bigcap_{l\epsilon A_j}\bigcap_{i=1}^{n+p}(V_{il} = v_{il})]. \tag{6.5}$$

This leads to an expression for t_{ji} given by (3.7) where a_{ji} is replaced by $a_{ji}+k_j$. Remember that the k_j's now are supposed to be known.

Regarding the consistency of the experts a necessary condition for t_{ji} to be nonnegative is still that the jth expert's probability assessments are consistent with the multivariate exponential distribution of the Marshall Olkin type. A necessary and sufficient condition for the t_{ji}'s to be nonnegative is now that:

$$\pi_{ji} = P[V_i > u | \bigcap_{l\epsilon A_j}\bigcap_{i=1}^{n+p}(V_{il} = v_{il})] \geq [b_{ji}/(b_{ji} + u)]^{a_{ji}}[b_{ji}/(b_{ji} + u)]^{k_j}.$$

This means that the survival probability beyond u from the ith shock may now be reduced by a multiplicative factor $[b_{ji}/(b_{ji} + u)]^{k_j}$ after the assessments of the jth expert compared to the à priori assessments. This is intuitively obvious since we now can observe failures. Note that this factor is decreasing in k_j, the number of failures the jth expert has access to and is increasing in b_{ji}. The factor is close to zero for the vague gamma distribution which can be applied as in Huseby (1988), and does not lead to inconsistencies.

Note that we have no ordering of the v_{il}'s in the background information (6.1) as opposed to the one in (2.3), see (2.2). Hence the claim of the leftmost inequality in (4.1) now vanishes.

The predictive probability (6.3) may now be established from the t_{ji}'s exactly as in Section 5. In the case where the background information is given by (6.2) instead of (6.1), the sets A_j, B_g, C_j, C_j' and G_g, and the integers q, n_g and k_j must be indexed for each $i = 1, \cdots, n+p$.

However, an advantage of the present approach is that the proper predictive probability can be arrived at in a fully Bayesian fashion as in Huseby (1988). This will come out as a special case of the deductions in the next section where the predictive probability is updated due to getting real data.

7. Updated predictions based on real data

In the present paper we have not yet considered how the predictive probability can be updated when getting real data. When these data represent survivals of components, or more generally are given both for individual and common shocks, this is straightforward. Introduce ($i = 1, \cdots, n+p$):

$$\begin{aligned} T_i &= \text{total time on test relative to the ith shock} \\ d_i &= \text{number of shocks of type } i \end{aligned} \tag{7.1}$$

By the same argument leading to (2.5) the updated predictive probability due to the data (7.1) equals:

$$\prod_{i=1}^{n+p} \left(\frac{b_i + t_i + T_i}{b_i + t_i + T_i + v_i}\right)^{a_i + d_i}. \tag{7.2}$$

The corresponding generalization of (6.3) has the exponent $a_i + d_i + m$ instead of $a_i + d_i$.

The proper Bayesian predictive probability based on (6.1) is deducted in the following:

$$P[\bigcap_{i=1}^{n}(Z_i > z_i)|t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n+p \bigcap (T_i, d_i), i = 1, \cdots, n+p]$$

$$= c_1 \int_{\theta_1=0}^{\infty} \cdots \int_{\theta_{n+p}=0}^{\infty} P[\bigcap_{i=1}^{n+p}(V_i > v_i)|\theta_1, \cdots, \theta_{n+p}]$$

$$\times p[t_{ji}, j = 1, \cdots k; i = 1, \cdots, n+p|\theta_1, \cdots, \theta_{n+p}]$$

$$\times p[(T_i, d_i), i = 1 \cdots, n+p|\theta_1, \cdots, \theta_{n+p}] \prod_{i=1}^{n+p} \frac{b_i^{a_i} \theta_i^{a_i - 1}}{\Gamma(a_i)} \exp(-b_i \theta_i) d\theta_1 \cdots d\theta_{n+p},$$

having applied the conditional independence of $\{t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n+p\}$ and $\{(T_i, d_i), i = 1, \cdots, n+p\}$ given $\theta_1, \cdots, \theta_{n+p}$. The expression above can be written as

$$\begin{aligned} &= c_2 \int_{\theta_1=0}^{\infty} \cdots \int_{\theta_{n+p}=0}^{\infty} p[t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n+p|\theta_1, \cdots, \theta_{n+p}] \\ &\times \prod_{i=1}^{n+p} [\theta_i^{a_i + d_i - 1} \exp(-(b_i + v_i + T_i)\theta_i)] d\theta_1 \cdots d\theta_{n+p}. \end{aligned} \tag{7.3}$$

To arrive at $p[t_{ji}, j=1,\cdots,k; i=1,\cdots,n{+}p|\theta_1,\cdots,\theta_{n+p}]$ we introduce the transformation of $(s_{1i},\cdots,s_{qi})$ into $(s_{1i},\cdots,s_{(q-k-1)i},t_{1i},\cdots,t_{ki},t_i)$. If this transformation is singular, we can choose another one instead. From (5.3) we establish $(i=1,\cdots,n{+}p, g=q-k,\cdots,q)$:

$$s_{gi} = f_{gi}(s_{1i},\cdots,s_{(q-k-1)i},t_{1i},\cdots,t_{ki},t_i), \tag{7.4}$$

where the f_{gi}'s are linear functions. Since the s_{gi}'s, $g=1,\cdots,q; i=1,\cdots,n+p$ are independent given $\theta_1,\cdots,\theta_{n+p}$ and gamma distributed with shape and scale parameters respectively equal to n_g and θ_i, we get from (7.4):

$$\begin{aligned}
&p[t_{ji}, j=1,\cdots,k; i=1,\cdots,n+p|\theta_1,\cdots,\theta_{n+p}] \\
&= c_3 \prod_{i=1}^{n+p} \int_{a_{(i)}}^{b_{(i)}} \int_{S_{(q-k-1)i}} \cdots \int \prod_{g=1}^{q-k-1} (s_{gi}^{n_g-1}) \\
&\times \prod_{g=q-k}^{q} [f_{gi}(s_{1i},\cdots,s_{(q-k-1)i},t_{1i},\cdots,t_{ki},t_i]^{n_g-1} \\
&\times \theta_i^m \exp(-\theta_i t_i) ds_{1i}\cdots ds_{(q-k-1)i}dt_i.
\end{aligned} \tag{7.5}$$

Here $S_{(q-k-1)i}$ is an area of integration and $a_{(i)}$ and $b_{(i)}$ integration bounds determined by (5.3). Inserting (7.5) into (7.3) we finally get:

$$\begin{aligned}
&P[\bigcap_{i=1}^{n}(Z_i > z_i)|t_{ji}, j=1,\cdots,k; i=1,\cdots,n+p\bigcap(T_i,d_i), i=1,\cdots,n+p] \\
&= c\prod_{i=1}^{n+p} \int_{a_{(i)}}^{b_{(i)}} (b_i+t_i+T_i+v_i)^{-(a_i+d_i+m)} \int_{S_{(q-k-1)i}} \cdots \int \prod_{g=1}^{q-k-1} (s_{gi}^{n_g-1}) \\
&\times \prod_{g=q-k}^{q} [f_{gi}(s_{1i},\cdots,s_{(q-k-1)i},t_{1i},\cdots,t_{ki},t_i)]^{n_g-1} ds_{1i}\cdots ds_{(q-k-1)i}dt_i.
\end{aligned} \tag{7.6}$$

Here c is a normalizing constant ensuring that $v_i=0, i=1,\cdots,n+p$ give a joint survival probability of 1.

By specializing $(T_i,d_i)=(0,0), i=1,\cdots,n+p$, we end up with the predictive probability promised at the end of the previous section. By furthermore considering the special case (5.4) we end up with the following generalized version of (3.15) in Huseby (1988):

$$\begin{aligned}
&P[\bigcap_{i=1}^{n}(Z_i > z_i)|t_{ji}, j=1,\cdots,k; i=1,\cdots,n+p] \\
&= c\prod_{i=1}^{n+p} \int_{a_{(i)}}^{b_{(i)}} (b_i+t_i+v_i)^{-(a_i+m)} (t_i^* - t_i/(k-1))^{n_{(k+1)}-1} \\
&\times \prod_{j=1}^{k} [(t_{ji}-t_i^*) + t_i/(k-1)]^{n_j-1} dt_i,
\end{aligned} \tag{7.7}$$

where

$$t_i^* = \sum_{j=1}^{k} t_{ji}/(k-1)$$

and applying (5.5) and (5.6)

$$a_{(i)} = \sum_{j=1}^{k} t_{ji} - (k-1)\min\{t_{ji}|j = 1, \cdots, k\}$$

$$b_{(i)} = \sum_{j=1}^{k} t_{ji}.$$

When the data include failures of components, without knowing which shocks that occured, updating is far more difficult due to the intractability of the density function of the multivariate exponential distribution of the Marshall-Olkin type. This is not discussed in the present paper.

8. Some final comments

It should finally be noted that the use of expert opinions is actually implemented in the regulatory work for nuclear power plants in the US. In addition there is no reason that this approach should not be used for instance in the offshore oil industry when new or redesigned systems are analysed. A general problem when using expert opinions is the selection of the experts. This problem is not addressed directly in the present paper except for suggesting in Section 4 when a selected expert should be dismissed. However, asking experts technical questions on the component level as in the present paper, where the consequences for the overall reliability assessment on the system level are less clear, seems very advantageous. Hence one can avoid the problem that to any risk assessment on system level there is an expert that will strongly support it.

Acknowledgement

The present paper is following up research carried out in the project "Failure Data Information with Bayesian Statistics". This project was carried out inside the Nordic cooperative program Terotechnology, supported by the Royal Norwegian Council for Scientific and Industrial Research and the Nordic Industrial Fund. The authors are grateful to these institutions for making this research possible.

References

Barlow, R.E. and Proschan, F. (1975) *Statistical Theory of Reliability and Life Testing. Probability Models.* Holt, Rinehart and Winston, New York.

Christensen, P. and Kongsö, H.E. (1991) The use of on-line reliability analysis for maintenance planning. In *Operational Reliability and Systematic Maintenance*, ed. A. Folkeson & K. Holmberg. Elsevier, pp.

Huseby, A.B. (1986) Combining experts' opinions, a retrospective approach. Tech. Rep. Center for Industrial Research, P.O. Box 350, Blindern, Oslo 3, Norway.

Huseby, A.B. (1988) Combining opinions in a predictive case. *In Bayesian Statistics* **3**, ed. J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, Oxford University Press, pp. 641-651.

Lindley, D.V. and Singpurwalla, N.D. (1986) Reliability (and fault tree) analysis using expert opinions. *J. Amer. Statist. Ass.*, **81**, 87-90.

Natvig, B. and Eide, H. (1987) Bayesian estimation of system reliability. *Scand. J. Statist.*, **14**, 319-327.

Natvig, B. (1990) Using expert opinions in Bayesian estimation of system reliability. Submitted to the Proceedings of the Course-Congress on Reliability and Decision Making, Siena, Italy, October 15-26, 1990.

Tierney, L. and Kadane, J.B. (1986) Accurate approximations for posterior moments and marginal densities. *J. Amer. Statist. Ass.*, **81**, 82-86.

RÉSUMÉ:

Dans cet article, nous considérons le problème de combiner l'opinion de k experts sur les durées de vie des n composants d'un système binaire. Ce problème a été étudié dans le cas d'un seul composant par Huseby (1986, 1988). Comme les experts utilisent souvent certaines sources d'information communes, il explique que leurs estimations seront typiquement dépendantes, et que cette difficulté ne peut être surmontée sans émettre des jugements sur les sources d'information présentes et sur leurs accéssabilités à chaque expert. Dans le premier article, l'information accessible aux experts est modelée comme un ensemble d'observations $Y_1, \cdots, Y_m$. Ces observations sont alors reconstruites autant que possible d'après l'information fournie par les experts, et utilisées comme base du jugement combiné. Cette méthode est appelée l'approche rétrospective. Dans le deuxième article, la quantité inconnue est modelée comme une observation future Y, de même distribution que les Y_i. Cette méthode est appelée l'approche prédictive. Dans le cas $n > 1$, où chaque expert émet des opinions sur plus qu'un des composants, des dépendances additionelles entrent en jeu entre les fiabilités des composants. Ceci est par exemple vrai si au moins deux des composants sont soit de type similaire, soit partagent un environment commun ou sont exposés à des défectuosités de source commune. Dans le cas $n = 2$, les méthodes rétrospective et prédictive sont toutes deux considérées par Natvig (1990). Dans l'article présent, nous étudions l'approche prédictive dans le cas d'un nombre arbitraire n et d'une superposition arbitraire des ensembles d'observation des différents experts. Les durées de vie des composants sont supposées avoir une distribution exponentielle multivariée du type de Marshall-Olkin. A la fin de l'article, nous montrons comment la distribution simultanée des durées de vie des n composants peut facilement être remise à jour si l'on recoit des données véridiques.