Expert Opinions in Bayesian Estimation of System Reliability in a Shock Model - the MTP_2 Connection

JØRUND GÅSEMYR and BENT NATVIG
University of Oslo

ABSTRACT. In this paper combining the opinions of k experts about the reliabilities of n components of a binary system is considered. This problem has been treated in the single component case by Huseby (1986, 1988) considering respectively the socalled retrospective and predictive approaches. For the two component case both approaches are treated in Natvig (1992), whereas the predictive approach is considered in Gåsemyr & Natvig (1991) 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. The present paper parallels the latter one considering the retrospective approach and also allowing for noisy assessments of the experts. We now arrive at the joint prior distribution of the reliabilities of the components. When this is MTP_2 (Multivariate Totally Positive of Order 2), it is shown that the machinary of Natvig & Eide (1987) can be applied to arrive at the posterior distribution of system reliability, based on data both on the component and system level. Hence a key question to be answered in the present paper is the following. When does the joint prior distribution of the reliabilities based on expert opinions in fact possess the MTP_2 property?

Key words: Multivariate exponential distribution, common cause failures, combined judgement, noisy assessments.

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 & 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., it was assumed that the components have independent prior reliabilities. $\pi_i(p_i)$ describes the 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) on the underlying joint distribution of the parameters in the model. 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 $n = 2$ both approaches are treated in Natvig (1992), whereas the predictive approach is considered in Gåsemyr & Natvig (1991) for an arbitrary n and 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. The present paper parallels the latter one considering the retrospective approach and also allowing for noisy assessments of the experts. From the assessments of the experts we arrive at the joint prior distribution, $\pi(\underline{p})$, of the reliabilities.

Let us now first consider the case of independent components given $\underline{p}$. Suppose that we run experiments on the component level and get the data $\underline{D} = (D_1, \cdots, D_n)$ where D_i is the data from the experiment on the ith component. Let $\pi(\underline{D}|\underline{p})$ be the corresponding likelihood function. Hence the posterior distribution of the reliabilities, $\pi(\underline{p}|\underline{D})$, is given by:

$$\pi(\underline{p}|\underline{D}) = \frac{\pi(\underline{D}|\underline{p})\pi(\underline{p})}{\int \pi(\underline{D}|\underline{p})\pi(\underline{p})d\underline{p}}. \tag{1.2}$$

The corresponding distribution of system reliability $\pi\{h(\underline{p})|\underline{D}\}$ can in *principle* be arrived at by using the transformation formula for joint probability distributions. The prior dependencies between $p_1, \cdots, p_n$ are not creating too much extra trouble here. By now using expert opinion on the system level, in the spirit of Huseby (1986), $\pi\{h(\underline{p})|\underline{D}\}$ may be updated to the prior distribution of system reliability $\pi_0\{h(\underline{p})|\underline{D}\}$. If we now finally run an experiment on the system level and get the data D, we end up with the posterior distribution of system reliability $\pi\{h(\underline{p})|\underline{D}, D\}$.

Let us next consider the case of associated components given $\underline{p}$. This is the challenging case modeling the nonnegative dependence between component states in real life systems. Due to this general assumption of dependence there is no way of establishing an exact expression for $\pi(h|\underline{D})$. It is not even possible to arrive at exact expressions for its first r moments. The best one can do is to arrive at bounds on these moments. From (1.2) the marginal posterior distribution of p_i, $\pi(p_i|\underline{D})$, is given by:

$$\pi(p_i|\underline{D}) = \frac{\int \pi(\underline{D}|\underline{p})\pi(\underline{p})d(\cdot_i,\underline{p})}{\int \pi(\underline{D}|\underline{p})\pi(\underline{p})d\underline{p}},$$ (1.3)

where $(\cdot_i,\underline{p}) = (p_1,\cdots,p_{i-1},\cdot,p_{i+1},\cdots,p_n)$. This leads to the moments up till order $r(i = 1,\cdots,n; j = 1,\cdots,r)$:

$$E(p_i^j|\underline{D}) = \frac{\int p_i^j\pi(\underline{D}|\underline{p})\pi(\underline{p})d\underline{p}}{\int \pi(\underline{D}|\underline{p})\pi(\underline{p})d\underline{p}},$$ (1.4)

by for instance applying an approximation technique suggested by Laplace, which has been pointed out to be quite good by Tierney & Kadane (1986).

From (1.4) by applying results of Natvig & Eide (1987) and of a very recent paper Lindqvist (1991) we arrive at bounds on:

$$E(h^j|\underline{D}), \quad j = 1,\cdots,r,$$ (1.5)

of $\pi(h|\underline{D})$.

However, the best bounds in these papers are based on the assumption that $p_1,\cdots,p_n$ are independent given $\underline{D}$. Sufficient conditions for this are that the components have independent prior reliabilities, which is unrealistic when the opinions of experts are used, and that $D_1,\cdots,D_n$ are independent given $\underline{p}$, which is reasonable if for instance different laboratories are used for different components. From (1.5) one may adjust a proper $\pi(h|\underline{D})$, which may be further updated to $\pi_0(h|\underline{D})$ and $\pi(h|\underline{D},D)$ as in the case of independent components.

The rather good lower bounds of Theorem 2.7 of Natvig & Eide (1987) and some good upper bounds of Lindqvist (1991) are valid also under the weaker assumption that $p_1,\cdots,p_n$ are associated given $\underline{D}$. According to Theorem 4.2 of Karlin & Rinott (1980) the MTP_2 (Multivariate Totally Positive of Order 2) property is stronger than the property of association. From (1.2) and Proposition 3.3 of the latter paper (Property 2 of the Appendix) a sufficient condition for this weaker assumption to be true is that $\pi(\underline{D}|\underline{p})$ and $\pi(\underline{p})$ both are MTP_2.

Recall that a random vector $(Z_1,\cdots,Z_n)$ is MTP_2 if and only if its density, $f(\underline{z})$, is MTP_2, i.e., if:

$$f(\underline{x} \vee \underline{y})f(\underline{x} \wedge \underline{y}) \geq f(\underline{x})f(\underline{y}),$$

where for $\underline{x} = (x_1,\cdots,x_n), \underline{y} = (y_1,\cdots,y_n)$

$$\underline{x} \vee \underline{y} = (\max(x_1,y_1),\cdots,\max(x_n,y_n))$$
$$\underline{x} \wedge \underline{y} = (\min(x_1,y_1),\cdots,\min(x_n,y_n)).$$

For the time being we have no idea of how the MTP_2 property can be converted into assumptions on observable random quantities. In the case where $D_1, \cdots, D_n$ are independent given $\underline{p}$, we have:

$$\pi(\underline{D}|\underline{p}) = \prod_{i=1}^{n} \pi(D_i|p_i),\qquad(1.6)$$

and it follows from the latter proposition that $\pi(\underline{D}|\underline{p})$ is MTP_2. Hence a key question to be answered in the present paper is the following. When does $\pi(\underline{p})$ established by using expert opinions in fact possess the MTP_2 property?

The paper is organized in the following way. In Sections 2 and 3 we consider the case of a single common shock destroying all components. In the former section it is assumed that the assessments of the experts are without noise whereas the latter section includes noise. Discouraged by the technicalities faced in this section an alternative approach is presented in Section 4 that includes noisy assessments of the experts and which may be more easily generalized to cases with other shock structures. The main difference is that the observations on which the experts' assessments are based, now are modeled as independent life times for each component as well as times to occurence of common shocks. Some further indications of generalizations are given in Section 5. Results needed to establish the MTP_2 property are concentrated in the Appendix.

2. The single common shock case without noisy assessments

The component lifetimes are assumed to have a multivariate exponential distribution of the Marshall-Olkin type with as a start a single common shock destroying all components. The time until failure of the ith component due to an individual shock is denoted by $V_i, i = 1, \cdots, n$ whereas V_{n+1} is the time until the common shock occurs. The variables V_i are independent given the parameters $\theta_i, i = 1, \cdots, n+1$ and exponentially distributed. Here θ_i is the failure rate corresponding to V_i. Then the lifetime, Z_i, of the ith component satisfies $(i = 1, \cdots, n)$:

$$Z_i = \min\{V_i, V_{n+1}\}.$$

We now suppose that the background information of the experts, corresponding to their observation sets, is as in the main approach of Gåsemyr & Natvig (1991) in terms of m independent sets of survival times beyond specific time points for all components; i.e.,

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

Now define:

$$v_{il} = z_{il}, \quad i = 1, \cdots, n; l = 1, \cdots, m$$
$$v_{(n+1)l} = \max_{1 \leq i \leq n} z_{il}, \quad l = 1, \cdots, m\qquad(2.2)$$

The information (2.1) is clearly equivalent to:

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

4

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

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, having k_j elements, $j = 1, \cdots, k$. We have:

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

We assume that the prior distributions of $\theta_i, i = 1, \cdots, n+1$ 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, $j = 1, \cdots, k; i = 1, \cdots, n+1$.

Introduce:

$$t_{ji} = \sum_{l \in A_j} v_{il}, j = 1, \cdots, k; i = 1, \cdots, n+1$$

$$t_i = \sum_{l=1}^{m} v_{il}, i = 1, \cdots, n+1. \tag{2.4}$$

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. By Bayes' theorem the posterior distributions of $\theta_i, i = 1, \cdots, n+1$ both for the DM and the jth expert are independent gamma distributions with shape and scale parameters respectively equal to $(a_i, b_i + t_i)$ for the DM and $(a_{ji}, b_{ji} + t_{ji})$ for the jth expert, $j = 1, \cdots, k; i = 1, \cdots, n+1$.

The t_{ji}'s are arrived at in (3.8) of Gåsemyr & Natvig (1991) by asking the experts about the marginal and joint survival probabilities, beyond a certain time point u, for the components. The t_i's are estimated from the t_{ji}'s as in Section 5 of the latter paper.

Now the reliability of the ith component at time t is given by:

$$p_i = \exp[-(\theta_i + \theta_{n+1})t], i = 1, \cdots, n.$$

In order to find the joint distribution of $\underline{p}$, introduce:

$$\phi_i = \theta_i + \theta_{n+1}, i = 1, \cdots, n$$

$$\phi_{n+1} = \theta_{n+1}. \tag{2.5}$$

Hence:

$$\phi_i = -\ln p_i / t, i = 1, \cdots, n. \tag{2.6}$$

Since the transformation (2.5) is linear with constant Jacobian, the joint distribution of the ϕ_i's is given by the density:

$$c_1 \prod_{i=1}^{n} (\phi_i - \phi_{n+1})^{a_i-1} \exp[-(b_i + t_i)(\phi_i - \phi_{n+1})]$$

$$\times \phi_{n+1}^{a_{n+1}-1} \exp[-(b_{n+1} + t_{n+1})\phi_{n+1}]. \tag{2.7}$$

5

This density is clearly MTP_2 by Properties 2 and 5 of the Appendix. Integrating with respect to ϕ_{n+1} gives an MTP_2 density for $\phi_1, \cdots, \phi_n$ by Properties 1 and 4 of the Appendix since the area of integration is adjusted for by multiplying (2.7) by the indicator function:

$$I(\phi_{n+1} < \min_{1 \leq i \leq n} \phi_i)$$

Since the transformations in (2.6) are all decreasing in the p_i's and since the corresponding Jacobian matrix is diagonal, we can finally conclude by Properties 2 and 3 that the joint distribution of p given $t_i, i = 1, \cdots, n+1$ is MTP_2. By substituting (2.6) and $\phi_{n+1} = -\ln p_{n+1}/t$ in (2.7) its density is given by:

$$c_2 \int_{\substack{\max p_i \\ 1 \leq i \leq n}}^{1} \prod_{i=1}^{n} p_i^{(b_i+t_i)/t-1} (\ln(p_{n+1}/p_i))^{a_i-1}$$

$$\times p_{n+1}^{[b_{n+1}+t_{n+1}-\sum_{i=1}^{n}(b_i+t_i)]/t-1} (\ln p_{n+1})^{a_{n+1}-1} dp_{n+1}. \tag{2.8}$$

As in Gåsemyr & Natvig (1991) an alternative approach is possible in the case where $k_j, j = 1, \cdots, k$ are known. We then replace the background information (2.3) by:

$$\bigcap_{i=1}^{n+1} (V_{il} = v_{il}), l = 1, \cdots, m. \tag{2.9}$$

Hence we now have data on the times to shocks instead of survival times of the components. The deductions above are still valid by replacing the shape parameters a_i and a_{ji} in the posterior distributions of θ_i by $a_i + m$ and $a_{ji} + k_j$ respectively for $i = 1, \cdots, n+1$.

One advantage of this approach is that the modified version of (2.8) may be updated when getting real data. When these data represent survivals of components, or more generally are given both for individual shocks and the common shock, this is straightforward. Introduce $(i = 1, \cdots, n+1)$:

$$T_i = \text{total time on test relative to the ith shock}$$
$$d_i = \text{number of shocks of type } i. \tag{2.10}$$

The updated joint distribution of p given $t_i, i = 1, \cdots, n+1$ for the alternative approach is then given by (2.8) with a_i and b_i replaced by respectively $a_i + d_i + m$ and $b_i + T_i, i = 1, \cdots, n+1$.

However, the main advantage of the alternativ approach is that the proper joint distribution of p can be arrived at in a fully Bayesian fashion parallel to the deductions in Huseby (1988) and Gåsemyr & Natvig (1991) avoiding the estimation of the t_i's from the t_{ij}'s.

We now have the following updated joint probability density function of $\theta_i, i = 1, \cdots, n+1$:

$$g[\theta_1, \cdots, \theta_{n+1}|t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n+1 \cap (T_i, d_i), i = 1, \cdots, n+1]$$
$$= c_3 p[t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n+1|\theta_1, \cdots, \theta_{n+1}]$$
$$\times p[(T_i, d_i), i = 1, \cdots, n+1|\theta_1, \cdots, \theta_{n+1}] \prod_{i=1}^{n+1} \frac{b_i^{a_i} \theta_i^{a_i-1}}{\Gamma(a_i)} \exp(-b_i \theta_i),$$

6

having applied the conditional independence of $\{t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n + 1\}$ and $\{(T_i, d_i), i = 1, \cdots, n + 1\}$ given $\theta_1, \cdots, \theta_{n+1}$. The first contribution above is given by (7.5) in Gåsemyr & Natvig (1991) and we end up with:

$$g[\theta_1, \cdots, \theta_{n+1} | t_{ji}, j = 1, \cdots, k; i = 1, \cdots, n + 1 \cap (T_i, d_i), i = 1, \cdots, n + 1]$$

$$= c_4 \prod_{i=1}^{n+1} \{\theta_i^{\alpha_i + d_i + m - 1} \exp(-(b_i + T_i)\theta_i) \int_{a_{(i)}}^{b_{(i)}} \exp(-\theta_i t_i)[\int_{S_{(q-k-1)i}} \cdots \int \prod_{g=1}^{q-k-1} (s_{gi}^{n_g-1})$$

$$\times \prod_{g=q-k}^{q} (f_{gi}(s_{1i}, \cdots, s_{(q-k-1)i}, t_{1i}, \cdots, t_{ki}, t_i))^{n_g-1} ds_{1i} \cdots ds_{(q-k-1)i}] dt_i\}. \tag{2.11}$$

Here we have introduced a disjoint partition B_g, having n_g elements, $g = 1, \cdots, q$ 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:

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

Furthermore:

$$s_{gi} = \sum_{k \in B_g} v_{il}, g = 1, \cdots, q; i = 1, \cdots, n + 1.$$

Then from (2.4), which is still valid for the alternative approach, we have:

$$t_{ji} = \sum_{g \in C_j} s_{gi}, j = 1, \cdots, k; i = 1, \cdots, n + 1$$

$$t_i = \sum_{g=1}^{q} s_{gi}, i = 1, \cdots, n + 1. \tag{2.12}$$

We then transform $(s_{1i}, \cdots, s_{qi})$ into $(s_{1i}, \cdots, s_{(q-k-1)i}, t_{1i}, \cdots, t_{ki}, t_i)$. From (2.12) we establish $(i = 1, \cdots, n + 1, g = q - k, \cdots, q)$:

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

where the f_{gi}'s are linear functions. (2.11) is arrived at by noting that the s_{gi}'s, $g = 1, \cdots, q, i = 1, \cdots, n + 1$ are independent given $\theta_1, \cdots, \theta_{n+1}$ and gamma distributed with shape and scale parameters respectively equal to n_g and θ_i. $S_{(q-k-1)i}$ is an area of integration and $a_{(i)}$ and $b_{(i)}$ integration bounds determined by (2.12).

The proper joint distribution of $\underline{p}$ is now arrived at completely parallel to the deduction of (2.8). The crucial question is whether this is MTP_2 as well. We are not able to prove this since by applying (2.5) we end up with a factor $\exp(-\phi_i t_i)$ in the integrand in (2.11). This is easily seen not to be MTP_2 in ϕ_i and t_i, which is necessary to apply Property 2 of the Appendix.

7

3. The single common shock case with noisy assessments of the experts

We return to the main approach of Section 2. However, now the experts' assessments are noisy, so instead of observing $Z_{il}, i = 1, \cdots, n; l \epsilon A_j, j = 1, \cdots, k$ they observe:

$$W_{jil}, i = 1, \cdots, n; l \epsilon A_j, j = 1, \cdots, k.$$

The distributions of $(W_{j1l}, \cdots, W_{jnl}), l \epsilon A_j, j = 1, \cdots, k$ are assumed to be independent each with a multivariate exponential distribution of the Marshall-Olkin type with parameters $\theta_1, \cdots, \theta_{n+1}$ as is the case for $(Z_{1l}, \cdots, Z_{nl}), l = 1, \cdots, m.$

When we say that the data are "observed" by the experts, as stated in Huseby (1986), we have in mind an intuitive process including a lot of subjective judgements and interpretations. Hence it may very well happen that the experts "observe" the data differently. Indeed when modeling experts' opinions it is difficult to say what is observation and what is interpretation. Introduce:

$$
\begin{aligned}
\underline{Z}_l &= (Z_{1l}, \cdots, Z_{nl}) \quad , l = 1, \cdots, m \\
\underline{Z} &= (\underline{Z}_1, \cdots, \underline{Z}_m) \\
Z_{(n+1)l} &= \max_{1 \le i \le n} Z_{il} \quad , l = 1, \cdots, m \\
T_{ji} &= \sum_{l \epsilon A_j} Z_{il} \quad , j = 1, \cdots, k; i = 1, \cdots, n+1 \\
\underline{T}_j &= (T_{j1}, \cdots, T_{j(n+1)}) \quad , j = 1, \cdots, k \\
W_{j(n+1)l} &= \max_{1 \le i \le n} W_{jil} \quad , l \epsilon A_j, j = 1, \cdots, k \\
W_{ji} &= \sum_{l \epsilon A_j} W_{jil} \quad , j = 1, \cdots, k; i = 1, \cdots, n+1 \\
\underline{W}_j &= (W_{j1}, \cdots, W_{j(n+1)}) \quad , j = 1, \cdots, k \\
\underline{W} &= (\underline{W}_1, \cdots, \underline{W}_k).
\end{aligned}
\tag{3.1}
$$

We assume that the DM assesses that the noisy assessments $\underline{W}_j, j = 1, \cdots, k$ from the different experts are independent given the "exact" ones $\underline{Z}$ and $\theta_1, \cdots, \theta_{n+1}$ with probability density functions on the form:

$$
g_j(\underline{w}_j | \underline{z}; \theta_1, \cdots, \theta_{n+1}) = g_j(\underline{w}_j | \underline{t}_j) = \prod_{i=1}^{n} g_{ji}(w_{ji} | t_{ji})
$$
$$
\times g_{j(n+1)}(w_{j(n+1)} | w_{j1}, \cdots, w_{jn}; t_{j1}, \cdots, t_{j(n+1)}).
\tag{3.2}
$$

Here $g_{ji}(w_{ji} | t_{ji})$ could be chosen with expectation $t_{ji}, i = 1, \cdots, n$ and $g_{j(n+1)}(w_{j(n+1)} | w_{j1}, \cdots, w_{jn}; t_{j1}, \cdots, t_{j(n+1)})$ with expectation $\max_{1 \le i \le n} w_{ji} + t_{j(n+1)} - \max_{1 \le i \le n} t_{ji}, j = 1, \cdots, k$. This would make the noisy total survival of the components centred close to the corresponding "exact" ones. Actually, as noted in Gåsemyr & Natvig (1991) there are uncertainties in the calculation of the t_{ij}'s and (3.2) is a way of modeling these uncertainties.

The posterior probability density function of $\theta_1, \cdots, \theta_{n+1}$ given the noisy assessments $\underline{w}$

from the experts is now arrived at in a fully Bayesian fashion:

$$h(\theta_1, \cdots, \theta_{n+1}|\underline{w})$$

$$= c_5 \int_{\underline{z}} \prod_{j=1}^{k} g_j(\underline{w}_j|\underline{z}; \theta_1, \cdots, \theta_{n+1}) f(\underline{z}|\theta_1, \cdots, \theta_{n+1}) \prod_{i=1}^{n+1} \frac{b_i^{\alpha_i} \theta_i^{\alpha_i - 1}}{\Gamma(\alpha_i)} \exp(-b_i \theta_i) d\underline{z} \qquad (3.3)$$

$$= c_6 \int_{\underline{z}} \prod_{j=1}^{k} g_j(w_j|\underline{t}_j) \prod_{l=1}^{m} f(\underline{z}_l|\theta_1, \cdots, \theta_{n+1}) \theta_i^{\alpha_i - 1} \exp(-b_i \theta_i) d\underline{z}.$$

Here $f(\underline{z}_l|\theta_1, \cdots, \theta_{n+1})$ is the complicated probability density function of the multivariate exponential distribution of the Marshall-Olkin type. From (3.3) one can as in the previous section derive the corresponding posterior joint distribution of $\underline{p}$ given the noisy assessments. This seems far from being MTP_2. If the deductions in this section had been based on the alternative approach from the previous one rather than the main approach, the joint distribution of $\underline{p}$ given the noisy assessments would have been simpler, but still not MTP_2.

4. A new alternative approach to cover noisy assessments of the experts

Discouraged by the technicalities we ran into in the previous section, we now discuss a new alternative approach to cover noisy assessments of the experts. This approach may be more easily generalized to cases with other shock structures than the one treated earlier in the present paper. The main difference is that the observations on which the experts' assessments are based, now are modeled as *independent* life times for each component as well as times to occurence of common shocks. Hence we have no longer a multivariate exponential distribution of the Marshall-Olkin type. The background information (2.9) is now replaced by:

$$\bigcap_{i=1}^{n+1}(Z_{il} = z_{il}) \quad , l = 1, \cdots, m. \qquad (4.1)$$

Here Z_{il} is the "exact" time to failure of the ith component, $i = 1, \cdots, n$, and $Z_{(n+1)l}$ the "exact" time until the common shock occurs in the lth observation set, $l = 1, \cdots, m$. The variables Z_{il} are assumed to be independent given the parameters $\phi_1, \cdots, \phi_{n+1}$ defined by (2.5) and exponentially distributed with failure rates $\phi_i, i = 1, \cdots, n+1; l = 1, \cdots, m$. The independence of the Z_{il}'s may seem unrealistic. 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.

The prior distributions of $\phi_i, i = 1, \cdots, n+1$ both for the DM and the jth expert are for convenience 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, $j = 1, \cdots, k; i = 1, \cdots, n+1$. Introduce:

$$t_{ji}^* = \sum_{l \in A_j} z_{il} \quad , j = 1, \cdots, k; i = 1, \cdots, n+1$$

$$t_i^* = \sum_{l=1}^{m} z_{il} \quad , i = 1, \cdots, n+1.$$

The t_{ji}^*'s are arrived at exactly as the t_{ji}'s, and the t_i^*'s are estimated from the t_{ji}^*'s exactly as the t_i's are estimated from the t_{ji}'s.

We now assume that the overall assessments of the experts, W_i^*, of $t_i^*, i = 1, \cdots, n+1$ are noisy. The variables W_i^* are assumed to be independent with expectation t_i^* given $t_i^*; \phi_i, i = 1, \cdots, n+1$ with conditional distributions given by:

$$h_i(w_i^*|t_1^*, \cdots, t_{n+1}^*; \phi_1, \cdots, \phi_{n+1}) = h_i(w_i^*|t_i^*), i = 1, \cdots, n+1.$$

The posterior probability density function of $\phi_1, \cdots, \phi_{n+1}$ given the noisy assessments $w_1^*, \cdots, w_{n+1}^*$ of the experts is now given by:

$$h(\phi_1, \cdots, \phi_{n+1}|w_1^*, \cdots, w_{n+1}^*)$$
$$= c_7 \prod_{i=1}^{n+1} \int_0^\infty h_i(w_i^*|t_i^*)t_i^{*(m-1)}\phi_i^{a_i^*-1}\exp(-(b_i^* + t_i^*)\phi_i)dt_i^*. \tag{4.2}$$

This density is MTP_2 by Property 2 of the Appendix since a single random variable is obviously MTP_2. Hence it follows as in Section 2 that the joint distribution of $\underline{p}$ given $w_i^*, i = 1, \cdots, n+1$ is MTP_2.

5. Some generalizations

As already stated the new alternative approach of Section 4 may be generalized to other shock structures. Suppose there are p possible common shocks. Introduce $(i = 1, \cdots, n)$:

$E_i = $ the set of common shocks that destroys the ith component.

Assume that the rth common shock occurs with failure rate $\theta_{n+r}, r = 1, \cdots, p$, and define:

$$\phi_i = \theta_i + \sum_{r \in E_i} \theta_{n+r} \quad , \quad i = 1, \cdots, n$$
$$\phi_i = \theta_i \qquad\qquad , \quad i = n+1, \cdots, n+p. \tag{5.1}$$

Concerning the background information the "exact" times to failure of the ith component, $i = 1, \cdots, n$ and until the rth common shock occurs, $r = 1, \cdots, p$ are assumed to be independent given the parameters $\phi_1, \cdots, \phi_{n+p}$ and exponentially distributed with failure rates $\phi_i, i = 1, \cdots, n+p$. This leads to a slightly modified version of (4.2) with $n+1$ replaced by $n+p$. $t_i^*, i = 1, \cdots, n+p$ are estimated as in Gåsemyr & Natvig (1991).

The methods of Sections 2 and 3 may easily be adjusted to situations with nonoverlapping common shocks. Less trivially they can also be adjusted to situations with a hierarchical shock structure. As an illustration suppose there are two common shocks with failure rates θ_{n+1} and θ_{n+2}. Define $(r = 1, 2)$:

$D_r = $ the set of components destroyed by the rth common shock,

and suppose $D_2 \subset D_1$. Define:

$$F_0 = \{1, \cdots, n\} - D_1, \quad F_1 = D_1 - D_2, \quad F_2 = D_2,$$

and the following parameters:

$$\phi_i = \theta_i \quad , \ i \epsilon F_0 \cup \{n+1\}$$
$$\phi_i = \theta_i + \theta_{n+1} \quad , \ i \epsilon F_1 \cup \{n+2\} \tag{5.2}$$
$$\phi_i = \theta_i + \theta_{n+1} + \theta_{n+2} \quad , \ i \epsilon F_2.$$

The modified version of (2.7) is:

$$c_8 \prod_{i \epsilon F_0 \cup \{n+1\}} \phi_i^{\alpha_i - 1} \exp[-(b_i + t_i)\phi_i]$$
$$\times \prod_{i \epsilon F_1 \cup \{n+2\}} (\phi_i - \phi_{n+1})^{\alpha_i - 1} \exp[-(b_i + t_i)(\phi_i - \phi_{n+1})] \tag{5.3}$$
$$\times \prod_{i \epsilon F_2} (\phi_i - \phi_{n+2})^{\alpha_i - 1} \exp[-(b_i + t_i)(\phi_i - \phi_{n+2})].$$

This is an MTP_2 density by the same argument as for (2.7) leading to a joint distribution of $\underline{p}$ given $t_i, i = 1, \cdots, n+2$ being MTP_2 as for (2.8).

Appendix: MTP_2 - Some properties

Property 1 Let f be an MTP_2 function on $\prod_{i=1}^{n} \mathcal{X}$. Then:

$$\phi(x_1, \cdots, x_{n-1}) = \int_{\mathcal{X}_1} f(x_1, \cdots, x_n) dx_1$$

is MTP_2 on $\prod_{i=1}^{n-1} \mathcal{X}_i$.

This is a special case of Proposition 3.2 of Karlin & Rinott (1980).

Property 2 Let f and g be MTP_2 functions. Then fg is MTP_2.

This is just Proposition 3.3 of the latter paper, which was referred to in Section 1.

Property 3 If $f(\underline{x}), \underline{x} \epsilon \prod_{i=1}^{n} \mathcal{X}_i$ is MTP_2, and $\phi_1, \cdots, \phi_n$ are all nondecreasing (or all nonincreasing) functions on $\mathcal{X}_1, \cdots, \mathcal{X}_n$, respectively, then the function:

$$\psi(\underline{x}) = f(\phi_1(x_1), \cdots, \phi_n(x_n))$$

is MTP_2 on $\prod_{i=1}^{n} \mathcal{X}_i$.

This is Proposition 3.6 of Karlin & Rinott (1980).

Property 4 The indicator function $I(x_1 < \min_{2 \le i \le n} x_i)$ is MTP_2.

Proof. Suppose $x_i \le y_i, i = 1, \cdots, n$. For any subset A of $\{2, \cdots, n\}$ we must show that:

$$I(y_1 < \min_{2 \le i \le n} y_i) I(x_1 < \min_{2 \le i \le n} x_i)$$
$$\ge I(y_1 < \min_{i \epsilon A, j \epsilon A^c} \{x_i, y_j\}) I(x_1 < \min_{i \epsilon A^c, j \epsilon A} \{x_i, y_j\}).$$

If $\min_{2 \leq i \leq n} x_i$ is obtained for an $i \epsilon A$, then:

$$I(y_1 < \min_{i \epsilon A, j \epsilon A^c}\{x_i, y_j\}) = I(y_1 < \min_{2 \leq i \leq n} x_i).$$

Since this factor is exceeded by both factors on the left hand side of the inequality, this case is done. If on the other hand $\min_{2 \leq i \leq n} x_i$ is obtained for an $i \epsilon A^c$, then:

$$I(x_1 < \min_{i \epsilon A^c, j \epsilon A}\{x_i, y_j\}) = I(x_1 < \min_{2 \leq i \leq n} x_i).$$

Hence the inequality is now straightforward and the proof is completed.

Property 5 Let g be a positive and continuous function of one variable and define $f(x_1, x_2) = g(x_2 - x_1)$. Then f is MTP_2 if and only if $\log(g)$ is concave.

Proof. f is MTP_2 if and only if for $x_i \leq y_i, i = 1, 2$:

$$g(y_2 - y_1)g(x_2 - x_1) \geq g(y_2 - x_1)g(x_2 - y_1).$$

Since clearly:

$$x_2 - y_1 \leq y_2 - y_1 \leq y_2 - x_1 \quad \text{and} \quad x_2 - y_1 \leq x_2 - x_1 \leq y_2 - x_1,$$

this is equivalent to:
$$g(c)g(d) \geq g(a)g(b)$$

for $a \leq c \leq b, a \leq d \leq b$ and $a + b = c + d$. Now take the logarithm on both sides and note that the claims on a, b, c, d are equivalent to

$$c = \lambda a + (1 - \lambda)b, d = (1 - \lambda)a + \lambda b$$

for some λ in $[0, 1]$. Setting $\lambda = 1/2$ it follows that $\log g$ is midpoint concave ($c = d = (a + b)/2$) and hence concave since g is continuous. If on the other hand $\log g$ is concave we have:

$$\log g(a) + \log g(b) = [\lambda \log g(a) + (1 - \lambda) \log g(b)]$$
$$+ [(1 - \lambda) \log g(a) + \lambda \log g(b)] \leq \log g(c) + \log g(d),$$

and the proof is completed.

References

Gåsemyr, J. & Natvig, B. (1991). Using expert opinions in Bayesian estimation of component lifetimes in a shock model - a general predictive approach. Invited paper presented at the 48th Session of the International Statistical Institute, Cairo, September 9-17, 1991.

Huseby, A.B. (1986). Combining experts' opinions, a retrospective approach. Tech. Rep., Centre for Industrial Research, Oslo.

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

Karlin, S. & Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Analysis* 10, 467-498.

Lindqvist, B. (1991). The minimal path upper bound for the moments of a random reliability function. Submitted to *Scand. J. Statist.*

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

Natvig, B. (1992). Using expert opinions in Bayesian estimation of system reliability. In *Reliability and Decision Making* (ed. R.E. Barlow, C.A. Clarotti & F. Spizzichino), pp. , Elsevier.

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