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 so-called 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 machinery 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 ($i = 1, \cdots, n$):

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

$$X = (X_1, \cdots, X_n),$$

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

Let furthermore:

$$E(X_i|p_i) = p_i = \text{the reliability of the } i\text{th component},$$

$$E(\phi(X)|h) = h = \text{the reliability of the system}.$$ 

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

$$h = E(\phi(X)|p) = h(p).$$
Natvig & Eide (1987) assumed that the joint prior distribution of the reliabilities, before running any experiments on the component level, \( \pi(p) \), can be written as:

\[
\pi(p) = \prod_{i=1}^{n} \pi_i(p_i),
\]

(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 \( Y_1, \ldots, Y_m \). 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, \( \pi(p) \), of the reliabilities.

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

\[
\pi(p|D) = \frac{\pi(D|p)\pi(p)}{\int \pi(D|p)\pi(p)dp}.
\]

(1.2)

The corresponding distribution of system reliability \( \pi(h(p)|D) \) can in principle be arrived at by using the transformation formula for joint probability distributions. The prior dependencies between \( p_1, \ldots, p_n \) are not creating too much extra trouble here. By now using expert opinion on the system level, in the spirit of Huseby (1986), \( \pi(h(p)|D) \) may be updated to the prior distribution of system reliability \( \pi_0(h(p)|D) \). 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 \( \pi(h(p)|D, D) \).
Let us next consider the case of associated components given $p$. 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 $\pi(h|D)$. 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 $p$, $\pi(p|D)$, is given by:

$$\pi(p|D) = \frac{\int \pi(D|p)\pi(p)d(\cdot, p)}{\int \pi(D|p)\pi(p)dp}, \quad (1.3)$$

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

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

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:

$$E(h^j|D), \quad j = 1, \ldots, r, \quad (1.5)$$

of $\pi(h|D)$.

However, the best bounds in these papers are based on the assumption that $p_1, \ldots, p_n$ are independent given $D$. Sufficient conditions for this are that the components have independent prior reliabilities, which is unrealistic when the opinions of experts are used, and that $D_1, \ldots, D_n$ are independent given $p$, which is reasonable if for instance different laboratories are used for different components. From (1.5) one may adjust a proper $\pi(h|D)$, which may be further updated to $\pi(h|D, D)$ 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 $p_1, \ldots, p_n$ are associated given $D$. According to Theorem 4.2 of Karlin & Rinott (1980) the $\text{MTP}_2$ (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 $\pi(D|p)$ and $\pi(p)$ both are $\text{MTP}_2$.

Recall that a random vector $(Z_1, \ldots, Z_n)$ is $\text{MTP}_2$ if and only if its density, $f(\cdot)$, is $\text{MTP}_2$, i.e., if:

$$f(x \vee y)f(x \wedge y) \geq f(x)f(y),$$

where for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$

$$x \vee y = (\max(x_1, y_1), \ldots, \max(x_n, y_n))$$

$$x \wedge y = (\min(x_1, y_1), \ldots, \min(x_n, y_n)).$$
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 $D_1, \ldots, D_n$ are independent given $p$, we have:

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

and it follows from the latter proposition that $\pi(D|p)$ is $MTP_2$. Hence a key question to be answered in the present paper is the following. When does $\pi(p)$ 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 occurrence 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 $i$th component due to an individual shock is denoted by $V_i, i = 1, \ldots, n$ whereas $V_{n+1}$ is the time until the common shock occurs. The variables $V_i$ are independent given the parameters $\theta_i, i = 1, \ldots, n + 1$ and exponentially distributed. Here $\theta_i$ is the failure rate corresponding to $V_i$. Then the lifetime, $Z_i$, of the $i$th component satisfies ($i = 1, \ldots, n$):

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

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

$$\bigcap_{i=1}^{n} (Z_i > z_i), \quad l = 1, \ldots, m.$$  \hfill (2.1)

Now define:

$$v_u = z_i, \quad i = 1, \ldots, n; l = 1, \ldots, m$$

$$v_{(n+1)i} = \max_{1 \leq i \leq n} z_i, \quad l = 1, \ldots, m$$

(2.2)

The information (2.1) is clearly equivalent to:

$$\bigcap_{i=1}^{n+1} (V_u > v_u), l = 1, \ldots, m,$$  \hfill (2.3)
where the $V_w$'s are independent and exponentially distributed with failure rates, $\theta_i, i = 1, \ldots, n+1; l = 1, \ldots, m$.

The DM assesses that the $j$th expert has access to information on the $Z_l$'s in (2.1) or equivalently on the $V_l$'s in (2.3) for $l$ with indices in the set $A_j$, having $k_j$ elements, $j = 1, \ldots, k$. We have:

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

We assume that the prior distributions of $\theta_i, i = 1, \ldots, n+1$ both for the DM and the $j$th 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 $j$th expert, $j = 1, \ldots, k; i = 1, \ldots, n+1$.

Introduce:

$$t_{ji} = \sum_{\mu \in A_j} v_{\mu}, j = 1, \ldots, k; i = 1, \ldots, n+1$$

$$t_i = \sum_{l=1}^m v_l, i = 1, \ldots, n+1. \quad \text{(2.4)}$$

Here $t_{ji}$ is the total survival of components from the $i$th shock, corresponding to the information from the $j$th expert. $t_i$ is similarly the total survival corresponding to the whole set of information. By Bayes' theorem the posterior distributions of $\theta_i, i = 1, \ldots, n+1$ both for the DM and the $j$th 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 $j$th expert, $j = 1, \ldots, k; i = 1, \ldots, n+1$.

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 $i$th component at time $t$ is given by:

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

In order to find the joint distribution of $p_i$, introduce:

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

$$\phi_{n+1} = \theta_{n+1}. \quad \text{(2.5)}$$

Hence:

$$\phi_i = -\ln p_i/t, i = 1, \ldots, n. \quad \text{(2.6)}$$

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

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

$$\times \phi_{n+1}^{m+1-1} \exp[-(b_{n+1} + t_{n+1})\phi_{n+1}], \quad \text{(2.7)}$$

5
This density is clearly $MTP_2$ by Properties 2 and 5 of the Appendix. Integrating with respect to $\phi_{n+1}$ gives an $MTP_2$ density for $\phi_1, \ldots, \phi_n$ 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 $t_i, i = 1, \ldots, n + 1$ is $MTP_2$. By substituting (2.6) and $\phi_{n+1} = -\ln p_{n+1}/t$ in (2.7) its density is given by:

$$c_2 \int_{\max_{1 \leq i \leq n} p_i}^{1} \prod_{i=1}^{n} p_i^{(k_i+4)/t-1}(\ln(p_{n+1}/p_i))^{a_i-1}$$

$$\times [b_{n+1} + t_{n+1} - \sum_{i=1}^{n} (k_i+4)/(t-1)] p_{n+1} (\ln p_{n+1})^{a_{n+1}-1} dp_{n+1}. \tag{2.8}$$

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

$$\bigcap_{i=1}^{n+1} (V_d = v_d), l = 1, \ldots, m. \tag{2.9}$$

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 $a_{j_i}$ in the posterior distributions of $\theta_i$ by $a_i + m$ and $a_{j_i} + k_j$ respectively for $i = 1, \ldots, n + 1$.

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 ($i = 1, \ldots, n + 1$):

$$T_i = \text{total time on test relative to the } i\text{th shock}$$

$$d_i = \text{number of shocks of type } i. \tag{2.10}$$

The updated joint distribution of $p$ given $t_i, i = 1, \ldots, n + 1$ 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 $b_i + T_i, i = 1, \ldots, n + 1$.

However, the main advantage of the alternative 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 $t_j$'s.

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

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

$$= c_3 p[t_{j_i}, j = 1, \ldots, k; i = 1, \ldots, n + 1| \theta_1, \ldots, \theta_{n+1}]$$

$$\times p[T_i, d_i], i = 1, \ldots, n + 1| \theta_1, \ldots, \theta_{n+1}] \prod_{i=1}^{n+1} \frac{b_{i}^{a_i} \theta_i^{a_i-1}}{\Gamma(a_i)} \exp(-b_i \theta_i),$$
having applied the conditional independence of \{t_{ji}, j = 1, \ldots, k; i = 1, \ldots, n + 1\} and \{(T_i, d_i), i = 1, \ldots, n + 1\} given \(\theta_1, \ldots, \theta_{n+1}\). The first contribution above is given by (7.5) in Gåsemyr & Natvig (1991) and we end up with:

\[
g[\theta_1, \ldots, \theta_{n+1}|t_{ji}, j = 1, \ldots, k; i = 1, \ldots, n + 1 \cap (T_i, d_i), i = 1, \ldots, n + 1]
= c_4 \prod_{i=1}^{n+1} \{\phi_i^{a_i+4+m-1} \exp(-b_i T_i) \int \exp(-\theta_i t_i) \left[ \prod_{j=1}^{q-k-1} (s_{gj}^\gamma)^{(-)} \right] \}
\times \prod_{g=q-k}^q \left( f_{gi}(s_{ji}, s_{(q-k-1)i}, t_{1i}, \ldots, t_{ki}, t_i) \right) s_{ji}^{(-1)} ds_{ji} \cdots ds_{(q-k-1)i} dt_i. \tag{2.11}
\]

Here we have introduced a disjoint partition \(B_g\), having \(n_g\) elements, \(g = 1, \ldots, q\) of the set \(\{1, \ldots, m\}\) and subsets \(C_j, j = 1, \ldots, k\) of the set \(\{1, \ldots, q\}\) such that we have the representation:

\[A_j = \bigcup_{g \in C_j} B_g, \quad j = 1, \ldots, k.\]

Furthermore:

\[s_{gi} = \sum_{u \in B_g} u_i, g = 1, \ldots, q; i = 1, \ldots, n + 1.\]

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

\[t_{ji} = \sum_{g \in C_j} s_{gi}, j = 1, \ldots, k; i = 1, \ldots, n + 1 \tag{2.12}\]

\[t_i = \sum_{g=1}^q s_{gi}, i = 1, \ldots, n + 1.\]

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

\[s_{gi} = f_{gi}(s_{1i}, \ldots, s_{(q-k-1)i}, t_{1i}, \ldots, t_{ki}, t_i),\]

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

The proper joint distribution of \(p\) 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 \(\phi_i\) and \(t_i\), which is necessary to apply Property 2 of the Appendix.
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 \( Z_u, i = 1, \ldots, n; l \in A_j, j = 1, \ldots, k \) they observe:

\[
W_{ju}, i = 1, \ldots, n; l \in A_j, j = 1, \ldots, k.
\]

The distributions of \( (W_{j1}, \ldots, W_{jn}), l \in A_j, j = 1, \ldots, k \) are assumed to be independent each with a multivariate exponential distribution of the Marshall-Olkin type with parameters \( \theta_1, \ldots, \theta_{n+1} \) as is the case for \( (Z_{1i}, \ldots, Z_{ni}), l = 1, \ldots, m \).

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:

\[
Z_l = (Z_{1l}, \ldots, Z_{nl}), l = 1, \ldots, m
\]

\[
Z = (Z_1, \ldots, Z_m)
\]

\[
Z_{(n+1)i} = \max_{1 \leq i \leq n} Z_{ui}, l = 1, \ldots, m
\]

\[
T_{ji} = \sum_{l \in A_j} Z_{ui}, j = 1, \ldots, k; i = 1, \ldots, n + 1
\]

\[
T_j = (T_{j1}, \ldots, T_{j(n+1)}), j = 1, \ldots, k
\]

\[
W_{j(n+1)i} = \max_{1 \leq i \leq n} W_{ju}, l \in A_j, j = 1, \ldots, k
\]

\[
W_{ji} = \sum_{l \in A_j} W_{ju}, j = 1, \ldots, k; i = 1, \ldots, n + 1
\]

\[
W_j = (W_{j1}, \ldots, W_{j(n+1)}), j = 1, \ldots, k
\]

\[
W = (W_1, \ldots, W_k).
\]

We assume that the DM assesses that the noisy assessments \( W_j, j = 1, \ldots, k \) from the different experts are independent given the “exact” ones \( Z \) and \( \theta_1, \ldots, \theta_{n+1} \) with probability density functions on the form:

\[
g_j(w_j | z; \theta_1, \ldots, \theta_{n+1}) = g_j(w_j | t_j) = \prod_{i=1}^{n} g_j(w_{ji} | t_{ji})
\]

\[
\times g_j(n+1)(w_{j(n+1)} | w_{j1}, \ldots, w_{jn}, t_{j1}, \ldots, t_{j(n+1)}).
\]

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

The posterior probability density function of \( \theta_1, \ldots, \theta_{n+1} \) given the noisy assessments \( w \)
from the experts is now arrived at in a fully Bayesian fashion:

\[ h(\theta_1, \ldots, \theta_{n+1} | \mathbf{y}) = c_n \int \prod_{j=1}^k g_j(w_j | z; \theta_1, \ldots, \theta_{n+1}) f(z | \theta_1, \ldots, \theta_{n+1}) \prod_{l=1}^{n+1} \frac{b_l^{\alpha_l} \theta_l^{\alpha_l-1}}{\Gamma(\alpha_l)} \exp(-b_l \theta_l) \, dz \]

(3.3)

= c_n \int \prod_{j=1}^k g_j(w_j | t_j) \prod_{l=1}^m f(z_l | \theta_1, \ldots, \theta_{n+1}) \theta_l^{\alpha_l-1} \exp(-b_l \theta_l) \, dz_l.

Here \( f(z_l | \theta_1, \ldots, \theta_{n+1}) \) 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 \( \mathbf{p} \) 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 \( \mathbf{p} \) 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 occurrence 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:

\[ \bigcap_{i=1}^{n+1} (Z_l = z_l), \quad l = 1, \ldots, m. \]  

(4.1)

Here \( Z_l \) is the "exact" time to failure of the \( i \)-th component, \( i = 1, \ldots, n \), and \( Z_{n+1} \) the "exact" time until the common shock occurs in the \( l \)-th observation set, \( l = 1, \ldots, m \). The variables \( Z_l \) are assumed to be independent given the parameters \( \phi_1, \ldots, \phi_{n+1} \) defined by (2.5) and exponentially distributed with failure rates \( \phi_i, i = 1, \ldots, n+1; l = 1, \ldots, m \). The independence of the \( Z_l \)'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 \( \phi_i, i = 1, \ldots, n+1 \) both for the DM and the \( j \)-th expert are for convenience independent gamma distributions with shape and scale parameters respectively equal to \( (a^*_i, b^*_i) \) for the DM and \( (a^*_j, b^*_j) \) for the \( j \)-th expert, \( j = 1, \ldots, k; i = 1, \ldots, n+1 \). Introduce:

\[ t^*_j = \sum_{i=1}^{n+1} z_i, \quad j = 1, \ldots, k; \]

\[ t^*_i = \sum_{l=1}^m z_l, \quad i = 1, \ldots, n+1. \]
The $t^*_i$'s are arrived at exactly as the $t_i$'s, and the $t^*_i$'s are estimated from the $t^*_j$'s exactly as the $t_j$'s are estimated from the $t_i$'s.

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

$$h_i(w^*_i | t^*_1, \ldots, t^*_n, \phi_1, \ldots, \phi_{n+1}) = h_i(w^*_i | t^*_i), i = 1, \ldots, n + 1.$$  

The posterior probability density function of $\phi_1, \ldots, \phi_{n+1}$ given the noisy assessments $w^*_1, \ldots, w^*_{n+1}$ of the experts is now given by:

$$h(\phi_1, \ldots, \phi_{n+1} | w^*_1, \ldots, w^*_{n+1}) = c_1 \prod_{i=1}^{n+1} \int h_i(w^*_i | t^*_i) \phi^{n-1}_i \exp\left(-\left(b_i + t^*_i\right)\phi_i\right) dt^*_i. \quad (4.2)$$

This density is $MTP_3$ 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 $p$ given $w^*_i, i = 1, \ldots, n + 1$ 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 $(i = 1, \ldots, n)$:

$$E_i = \text{the set of common shocks that destroys the } i\text{th component.}$$

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

$$\phi_i = \theta_i + \sum_{r \in E_i} \theta_{n+r}, \quad i = 1, \ldots, n$$

$$\phi_i = \theta_i, \quad i = n + 1, \ldots, n + p. \quad (5.1)$$

Concerning the background information the "exact" times to failure of the $i$th component, $i = 1, \ldots, n$ and until the $r$th common shock occurs, $r = 1, \ldots, p$ are assumed to be independent given the parameters $\phi_1, \ldots, \phi_{n+p}$ and exponentially distributed with failure rates $\phi_i$, $i = 1, \ldots, n + p$. This leads to a slightly modified version of (4.2) with $n + 1$ replaced by $n + p$. $t^*_i, i = 1, \ldots, n + p$ are estimated as in Gásemýr & Natvíg (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 $\theta_{n+1}$ and $\theta_{n+2}$. Define $(r = 1, 2)$:

$$D_r = \text{the set of components destroyed by the } r\text{th common shock},$$

and suppose $D_2 \subset D_1$. Define:

$$F_0 = \{1, \ldots, n\} - D_1, \quad F_1 = D_1 - D_2, \quad F_2 = D_2.$$
and the following parameters:
\[
\begin{align*}
\phi_1 &= \theta_1, \quad \text{i.e.} F_0 \cup \{n + 1\} \\
\phi_2 &= \theta_1 + \theta_{n+1}, \quad \text{i.e.} F_1 \cup \{n + 2\} \\
\phi_3 &= \theta_1 + \theta_{n+1} + \theta_{n+2}, \quad \text{i.e.} F_2.
\end{align*}
\]

The modified version of (2.7) is:
\[
\begin{align*}
\alpha_k \prod_{i \in P_0 \cup \{n+1\}} \phi_i^{a_i-1} \exp[-(b_i + t_i)\phi_i] \\
\times \prod_{i \in P_1 \cup \{n+2\}} (\phi_i - \phi_{n+1})^{a_i-1} \exp[-(b_i + t_i)(\phi_i - \phi_{n+1})] \\
\times \prod_{i \in P_2} (\phi_i - \phi_{n+2})^{a_i-1} \exp[-(b_i + t_i)(\phi_i - \phi_{n+2})].
\end{align*}
\]

This is an \( MTP_2 \) density by the same argument as for (2.7) leading to a joint distribution of \( g \) 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 \( \prod_{i=1}^n X_i \). Then:
\[
\begin{align*}
\phi(x_1, \cdots, x_{n-1}) &= \int f(x_1, \cdots, x_n) \, dx_1 \\
\end{align*}
\]
is \( MTP_2 \) on \( \prod_{i=1}^{n-1} X_i \).

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 \( f(x), x \in \prod_{i=1}^n X_i \) is \( MTP_2 \), and \( \phi_1, \cdots, \phi_n \) are all nondecreasing (or all nonincreasing) functions on \( X_1, \cdots, X_n \), respectively, then the function:
\[
\psi(x) = f(\phi_1(x_1), \cdots, \phi_n(x_n))
\]
is \( MTP_2 \) on \( \prod_{i=1}^n X_i \).

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

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

Proof. Suppose \( x_i \leq y_i, i = 1, \cdots, n \). For any subset \( A \) of \( \{2, \cdots, n\} \) we must show that:
\[
\begin{align*}
I(y_1 < \min_{2 \leq i \leq n} y_i) I(x_1 < \min_{2 \leq i \leq n} x_i) \\
\geq I(y_1 < \min_{a_j \in A, j \in A} \{x_i, y_j\}) I(x_1 < \min_{a_j \in A, j \in A} \{x_i, y_j\}).
\end{align*}
\]
If $\min_{2 \leq i \leq n} x_i$ is obtained for an $i \in A$, then:

$$I(y_1 < \min_{i \in A \cup \{A'\}} \{x_i, y_i\}) = I(y_1 < \min_{2 \leq i \leq n} x_i).$$

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 $i \in A'$, then:

$$I(x_1 < \min_{i \in A \cup \{A'\}} \{x_i, y_i\}) = I(x_1 < \min_{2 \leq i \leq n} x_i).$$

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:

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

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

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

for some $\lambda$ 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:

$$\log(g(a) + \log(g(b)) = [\lambda \log(g(a)) + (1 - \lambda) \log(g(b))$$

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

and the proof is completed.

**References**


