Expert Judgment in Maintenance Optimization

Jan M. van Noortwijk
Delft University of Technology, Delft
Rommert Dekker, Member IEEE
Erasmus University, Rotterdam
Roger M. Cooke
Delft University of Technology, Delft
Thomas A. Mazzuchi, Member IEEE
The George Washington University, Washington DC

Key Words — **Histogram technique, Dirichlet distribution, Elicitation & combination of expert opinion, Supra Bayes approach**

Reader Aids —
Purpose: Widen state of the art
Special math needed for explanation: Probability, statistics
Special math needed to use results: Same
Results useful to: Reliability theoreticians, maintenance analysts

Summary & Conclusions — **This paper proposes a comprehensive method for the use of expert opinion for obtaining lifetime distributions required for maintenance optimization. The method includes procedures for the elicitation of discretized lifetime distributions from several experts, the combination of the elicited expert opinion into a consensus distribution, and the updating of the consensus distribution with failure and maintenance data. The method was motivated by the practical circumstances governing its implementation. In particular, by the lack of statistical training of the experts and the high demands on their time. The use of a discretized life distribution provides more flexibility, is more comprehendible by the experts in the elicitation stage, and greatly reduces the computation in the combination and updating stages. The methodology is Bayes, using the Dirichlet distribution as the prior distribution for the elicited discrete lifetime distribution. Methods are described for incorporating information concerning the expertise of the experts into the analysis.**

1. INTRODUCTION

In general, downtime of a production unit due to failure of one of its components, can induce high cost. Time-based preventive maintenance, consisting of inspections and replacements, can reduce the number of unit failures and thus lower this cost considerably; however, executing preventive maintenance activities too frequently can be costly as well. The aim of maintenance optimization is therefore to determine an optimal maintenance interval, such that the total mean cost of failures and preventive maintenance activities is minimal. Requirements for proper maintenance optimization are:

- a subdivision of a system or unit into components, with a specification of failure modes and a description of the effect of preventive maintenance activities on specific components
- a specification of cost figures for execution of a preventive maintenance activity and for component failure
- a specification of lifetime distributions of the components
- a comprehensive model with a (numerical) optimization routine.

Maintenance optimization has been of interest at Koninklijke/Shell-Laboratorium, Amsterdam, and has led to (among other things) the development of a decision support system for maintenance optimization, called PROMPT. From the experience gained with the PROMPT project, it was concluded that obtaining the component lifetime distributions is the bottleneck for implementation of maintenance optimization [1]. This is mainly due to the scarcity of reliable data for determining such a distribution. The reason is that, for properly maintained systems, failures are rarely observed. In fact, when too many failures are observed, the system may undergo modifications. In addition, available data are often incomplete, with actual component level information in only a limited number of cases. Thus the initialization of a decision support system for maintenance optimization must rely on the use of expert opinion for determining component lifetime distributions. The ability to update the life distribution estimates as actual data become available is also highly desirable. Thus a theoretically sound systematic procedure is required to address the questions:

- How can a lifetime distribution be elicited from an expert?
- How can the opinions of several experts be combined?
- How can the consensus distribution be updated with actual failure and maintenance data?

The use of expert opinion is not new; see Cooke [2], Mosleh, Bier, Apostolakis [3], Singpurwalla [4], Spetzler & Staël von Holstein [5], and Wallsten & Budescu [6]. But the development of a complete systematic procedure based on statistical principles for the use of expert opinion in a maintenance environment is new and presents a different set of constraints governing its application. We develop this procedure and give a brief overview of some of the considerations that had to be taken into account. For ease of notation we consider the procedure for a single component.

Notation

D	index for decision maker
E	number of experts
m	number of histogram (time) intervals for the component
n_{ie}	subjective number of failures in time-interval i specified by expert e for the component; $i=1,\dots,m$, and $e=1,\dots,E$
n	$\sum_{i=1}^{m} n_{ie}$ total subjective number of failures for a component in $(0,\infty)$

p_{ie} — n_{ie}/n, subjective probability of failure in interval i specified by expert e for the component; $i=1,\ldots,m$, and $e=1,\ldots,E$

$\boldsymbol{p}_e$ — $(p_{1e},\ldots,p_{me})$ subjective probability distribution of failure specified by expert e for the component; $e=1,\ldots,E$

$\boldsymbol{P}_A$ — $(P_{1A},\ldots,P_{mA})$ probability distribution of failure for the component

$\boldsymbol{p}_A$ — $(p_{1A},\ldots,p_{mA})$ realization of the random vector $\boldsymbol{P}_A$

$\boldsymbol{p}_D$ — $(p_{1D},\ldots,p_{mD})$ decision maker's (consensus) estimate of $\boldsymbol{P}_A$

r_i — number of recorded failures of the component in time-interval i used to calibrate the experts; $i=1,\ldots,m$

c_i — number of recorded maintenance activities for the component at time t_i used to calibrate the experts; $i=1,\ldots,m-1$

s_i — number of observed failures of the component in time-interval i used to update $\boldsymbol{p}_D$; $i=1,\ldots,m$

v_i — number of observed maintenance activities for the component at time t_i used to update $\boldsymbol{p}_D$; $i=1,\ldots,m-1$

Other, standard notation is given in "Information for Readers & Authors" at the rear of each issue.

2. ELICITATION

The selection of an appropriate procedure for eliciting expert subjective probabilities must take into account the nature of the information required, the background and training of the experts, the number of experts, and the number of variables to be assessed. In our problem we must determine the failure distribution of components which display aging (nonconstant) and revealed failure behavior. It is only by attempting to model this aging behavior that we can justify the use of a time-based preventive maintenance policy. Typically five or fewer experts are familiar with the *a* particular component. Most of the available experts have little knowledge or experience with statistical failure models, and have only limited time to spend with an analyst for the elicitation process. In one elicitation session several component lifetime distributions (in addition to other pertinent questions relating to the effects of preventive maintenance, the consequences of failure, etc.) must be assessed and so due to the time constraints and powers of concentration of the experts, the elicitation process must be kept as short as possible.

An extensive review of elicitation techniques is in Wallsten & Budescu [6] and Spetzler & Staël von Holstein [5]. None of the techniques discussed in these sources, however, was specifically designed for the present problem. The fact that entire life distribution is required, and not merely individual probability assessments, negates the use of so-called "indirect elicitation techniques", based on such methods as lotteries or paired comparisons. Even if a well-known failure model were assumed, it would have at least two parameters which could not be considered separately. Requiring experts to fit a parametric distribution to their subjective probabilities is a lot to ask even from statistically trained experts.

Many nontechnical people, however, are familiar with histograms; see Ibrekk & Morgan [7]. Indeed, experience in the PROMPT project with a simplified version of a histogram technique (asking only for failure probabilities from two different time intervals) proved quite acceptable to our experts [1]. Kabus [8] also reports that a histogram technique used to predict the interest rate of certificates of deposit performed very well in that the information obtained showed good predictive validity. In view of these experiences we decided to base our elicitation procedure on the histogram technique. In addition, from the practical point of view, it is more comprehendible for statistically untrained experts to use a discretised version of the continuous pdf (viz, a histogram) thus replacing the concept of probability density by the concept of the probability of failure in a fixed time interval. Moreover, combining and updating require far less computing effort in the discrete case than in the continuous case, and a continuous distribution can always be fit to the discrete probability function obtained.

3. THE HISTOGRAM TECHNIQUE

Let the domain of lifetimes $(0,\infty)$ be divided into m disjoint time intervals $(t_{i-1},t_i]$, $i=1,\ldots,m-1$, and (t_{m-1},t_m) where $0\equiv t_0 < t_1 <\ldots< t_{m-1} < t_m \equiv \infty$. The experts need only specify their subjective probability of failure for the component in each interval. For convenience we denote time interval $(t_{i-1},t_i]$ by time interval i, $i=1,\ldots,m-1$. Defining interval m, (t_{m-1},∞), as an open interval is motivated by the fact that maintenance engineers have experiences only with the first part of a component life cycle since most components are replaced before failing. Knowing exactly the tail of a lifetime distribution is often not possible and not necessary for maintenance optimization.

Although the histogram appears to be the best available technique for assessing the component lifetime distribution in our environment, several implementational issues must be addressed. Specifically the selection of t_i, m, and the method for obtaining the experts' interval failure probabilities. In considering each of these there is a tradeoff between limiting the accuracy obtained through elicitation and requiring an accuracy from the expert which cannot be achieved. Thus compromise solutions were sought and the emphasis was placed on making the elicitation process as easy for the expert as feasible.

Fitting the elicited discrete distribution to a continuous parametric family with two parameters, requires that at least three time intervals be specified. In addition, psychological experiments suggest that bias is reduced when choosing from maximally five alternatives [4]. Therefore we advise $m=5$.

The choice of $t_1,\ldots,t_{m-1}$ under the ordering restriction does not influence the developed mathematical model; however, for better visual perception of the failure behavior of the component, we recommend that these values be equidistant. For example, use $t_i=i^*x'$, $i=0,\ldots,m-1$, where x' represents a maintenance interval used in the past. This is advantageous since the maintenance expert's familiarity with x' enables the assessment of failure probabilities by a comparison of known maintenance intervals. For ease in combining the elicited distributions, we recommend letting an analyst select x' before elicitation begins. A disadvantage in the above approach is that the elicited distribution can have a large tail, ie, most components are anticipated to fail in the last time interval. This is only a problem when the maintenance optimization indicates that the optimal maintenance interval is also somewhere in the tail. If this is the case, and the expert does not wish to change any of his input then a continuous distribution may have to be fitted to the elicited probability function and a more detailed maintenance optimization can be performed.

In eliciting subjective probabilities, Spetzler & Staël von Holstein [5] found that indicting a probability of say 0.001 as "one in one thousand" yielded better performance in elicitation. Applying this concept, we ask the expert to image that there are n components of the same type installed at time t_0 and request the expert to provide his anticipated number, n_{ie}, of components which fail within time interval i. We can then calculate experts e's (subjective) probability of failure within time interval i as $p_{ie} = n_{ie}/n$, $i = 1,\ldots,m$. As most people are accustomed to dealing with percentages, we define $n = 100$ to make the process even easier for the expert.

In order to enhance the speed and accuracy of the elicitation, an interactive PC-based program was developed. Using only the cursor controls, the expert generates a histogram display by iteratively increasing (or decreasing) the number of time intervals and defining the (subjective) number of anticipated failures in each interval. The expert receives on-line visual feedback on his assessment. Reduction of possible ambiguity or confusion is achieved with a trained analyst guiding the expert through the use of the program.

4. COMBINING OPINIONS FROM SEVERAL EXPERTS

Numerous methods have been developed for reaching a consensus using expert opinion from several sources. An excellent guide through the literature on aggregating expert opinions is Genest & Zidek [9]. Other studies are French [10], Cooke [2], and Winkler [11]. We propose a Bayes scheme that arises out of two main methods of combination: the *weighting scheme* and the *supra Bayes approach*. Other aggregation methods, eg, Delphi and paired-comparisons, are not considered in view of their time-consuming nature.

Both the weighting scheme and the supra Bayes approaches involve determining weights w_e, $e = 1,\ldots,E$, for each expert and determining a combination rule. The two approaches differ in their orientation. In weighting schemes, interest centers around determining an optimal fashion in which to use weights for combining the experts' opinions, while in the supra Bayes approach the emphasis is on analyzing and incorporating the background information which leads to the expert opinion assessed. This includes evaluation of the experts' expertise, their prior information sets, and the expert's prediction ability. This information is captured in the specification of a prior distribution for the quantity of interest and in the determination of an appropriate likelihood function for the expert opinion. The actual combining process is not a problem in the supra Bayes approach as all available information is conveyed in the form of a posterior distribution which is obtained in a straightforward manner using Bayes theorem.

Supra Bayes approaches for discrete probability functions have been suggested by Lindley [12], Morris [13], Mendel & Sheridan [14], and Winkler [11]. In a series of articles introduced by Winkler [15] and in Kempthorne & Mendel [16], conceptual problems which arise by application of the supra Bayes approach of Morris [13] are discussed. The problem with the supra Bayes approach of Lindley [12] is mainly its dependence on the specification of the covariances between the experts. The Bayes model of Mendel & Sheridan [14] is suited for combining the elicited distributions of two experts and has a slow start, since many seed quantities (quantities for which the true values are known) have to be asked. Therefore, of the four Bayes approaches, only the approach of Winkler [11] is left.

4.1 Development of the Supra Bayes Approach

Winkler [11] considers the case where the prior distribution of interest is a *natural conjugate of* the likelihood function. A strong advantage of the natural conjugate approach is that the application of Bayes theorem involves only algebraic manipulations. We extend the Bernoulli example in Winkler [11] to the multinomial situation encountered when eliciting histograms. We define the following likelihood for each expert e,

$$\mathcal{L}_e(\boldsymbol{p}_e|\boldsymbol{p}_A) = \prod_{i=1}^{m} (p_{iA})^{w_e \beta n p_{ie}} = \prod_{i=1}^{m} (p_{iA})^{w_e \beta n_{ie}} \qquad (4\text{-}1)$$

where the weight $w_e \geq 0$, is determining from the prior beliefs about the individual expert, and β is determined from the prior beliefs about the experts as a group. More detailed information on the determination of these parameters is presented in the following sections. The exponent in (4-1) is *the number of virtual observations* because it plays the mathematical role of the number of observations in the likelihood function when analyzing empirical failure data. This usage contrasts slightly with the notion of "equivalent observations" as used by Winkler [11].

Assuming conditional s-independence, we obtain the likelihood for all available expert opinion as:

$$\mathcal{L}(\boldsymbol{p}_1,\ldots,\boldsymbol{p}_E|\boldsymbol{p}_A) = \prod_{e=1}^{E} \left\{ \prod_{i=1}^{m} (p_{iA})^{w_e \beta n_{ie}} \right\}$$

$$= \prod_{i=1}^{m} (p_{iA})^{\beta \sum_{e=1}^{E} w_e n_{ie}} \qquad (4\text{-}2)$$

The $\beta \sum_{e=1}^{E} w_e n_{ie}$ corresponds to the number of virtual observations attributed to the experts as a group and thus β controls its magnitude.

It is well known [17] that the natural conjugate family for the multinomial likelihood is the Dirichlet distribution:

$$\pi(p_{1A},\ldots,p_{m-1A}|\alpha) = \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{m}\Gamma(\alpha_i)}\prod_{i=1}^{m} p_{iA}^{\alpha_i-1} \qquad (4\text{-}3)$$

where $\alpha_i > 0$, $i=1,\ldots,m$, $\Sigma_{i=1}^{m}\alpha_i=\alpha_0$, $p_{mA}=1-\Sigma_{i=1}^{m-1}p_{iA}$. As (4-3) is defined over the simplex $\{p_{iA}, i=1,\ldots,m: p_{iA} > 0, i=1,\ldots,m$, and $\Sigma_{i=1}^{m} p_{iA}=1\}$, it is well suited as a prior distribution for $\boldsymbol{P}_A$. In addition, the decision maker can incorporate his prior opinion about the probability function in a straightforward and easy fashion using—

$$\mathrm{E}\{P_{iA}\} = \frac{\alpha_i}{\alpha_0}, \qquad i=1,\ldots,m$$

$$\mathrm{Var}\{P_{iA}\} = \frac{\alpha_i(\alpha_0-\alpha_i)}{\alpha_0^2(\alpha_0+1)}, \qquad i=1,\ldots,m. \qquad (4\text{-}4)$$

Thus if *a priori*, the decision maker's best guess about the true value of $\boldsymbol{P}_A$ is expressed via $\boldsymbol{p}_A^* = (p_{1A}^*,\ldots,p_{mA}^*)$, he may incorporate this into the prior distribution by setting $\alpha_i=p_{iA}^*\alpha_0$, $i=1,\ldots,m$. As the value α_0 controls the variance of the P_{iA}, the decision maker may express his strength of belief in his prior estimate of $\boldsymbol{P}_A$ through α_0, ie, large (small) values of α_0 correspond to a high (low) degree of belief in the prior estimates and thus a low (high) prior variance.

In our case, the decision maker is knowledgeable about the experts but not about the failure behavior of the component. To reflect this, we define the decision maker's prior distributions for $\boldsymbol{P}_A$ as given by (4-3) with all α_i (and thus α_0) extremely small thus indicating the decision maker's low belief in his estimates.

Given the likelihood of all expert opinion (4-2) and the prior distribution (4-3) with every α_i of a minimal value, the posterior distribution for $\boldsymbol{P}_A$ is again a Dirichlet distribution of the form of (4-3) with the new parameters $\alpha_i \approx \beta\,\Sigma_{e=1}^{E} w_e n_{ie}$, $i=1,\ldots,m$. The use of "$\approx$" is due to disregarding the negligible prior parameter values. The decision maker's (consensus) probability distribution, $\boldsymbol{p}_D$, is obtained as the posterior estimate of $\boldsymbol{P}_A$ (posterior to the experts' probability distributions), which is derived from the posterior means (4-4):

$$\mathrm{E}\{P_{iA}|\boldsymbol{p}_1,\ldots,\boldsymbol{p}_E\} = \frac{\beta\sum_{e=1}^{E} w_e n_{ie}}{\sum_{i=1}^{m}\beta\sum_{e=1}^{E} w_e n_{ie}}$$

$$= \frac{\sum_{e=1}^{E} w_e p_{ie}}{\sum_{e=1}^{E} w_e}, \; i=1,\ldots,m. \qquad (4\text{-}5)$$

Winkler [11] proposed the bounds $1 \le \Sigma_{e=1}^{E} w_e \le E$ for the expert weights where the lower (upper) bound is attained if the experts base their opinion on identical (different) information sets. Under these bounds, the decision maker's distribution using the supra Bayes approach is identical to that if the decision maker had used the *Linear Opinion Pool* in Stone [18] and thus the posterior estimate is proper probability distribution which possesses (for $m > 2$) the *Marginalization Property* and the *Zero Preservation Property* (McConway [19]). In our specific application we consider experts who share (approximately) the same information set and thus restrict $\Sigma_{e=1}^{E} w_e=1$ but the more general case is easily accommodated.

4.2 Determining Weights for the Experts

Several methods for determining weights have been proposed in the literature such as:

1. equal weights
2. weights proportional to ranking of experts in terms of expertise
3. weights proportional to experts' self-ratings [11]
4. peer weights in which the experts rate each other's expertise [20]
5. Bayes updated weights [21]
6. calibration-information weights [22].

All of these weights have drawbacks with respect to the problem at hand. The ratings used in #2-#4 are not given any operational meaning, hence the notion of a "fair" self-rating or "fair" peer rating is not defined. Winkler [23] has shown that #5 is highly improper in that an expert can maximize his anticipated weight by assigning probability one to the interval in which he thinks failure is most likely to occur. The #6 has an intuitive appeal and has proved useful in practice, but requires many observations, which are generally not available in our context. The #1 (equal weights) is the simplest and has an intuitive appeal, however, it might not represent an optimal use of expert judgment.

In determining an optimal set of weights, we look for a solution which yields the best distribution for the decision maker from a mathematical point of view. The use of the likelihood of existing failure and maintenance data, given the consensus probability distribution $p_{iD} = \Sigma_{e=1}^{E} w_e p_{ie}$, $i=1,\ldots,m$, provides a useful measure even with only a limited amount of existing data. Specifically let $\boldsymbol{r} \equiv (r_1,\ldots,r_m)$, $\boldsymbol{c} \equiv (c_1,\ldots,c_m)$, and $c_m \equiv 0$; the likelihood function is:

$$\mathcal{L}(\boldsymbol{r}, \boldsymbol{c}|\boldsymbol{p}_D) = \prod_{i=1}^{m}(p_{iD})^{r_i}\left(\sum_{j=i+1}^{m} p_{jD}\right)^{c_i}. \qquad (4\text{-}6)$$

The weights can then be determined using:

$$\boldsymbol{w} = \underset{\boldsymbol{w}}{\mathrm{ARGMAX}}\left\{\mathcal{L}\left(\boldsymbol{r},\boldsymbol{c}\,\Big|\,\sum_{e=1}^{E} w_e\boldsymbol{p}_e\right)\right\}$$

$$\sum_{e=1}^{E} w_e=1. \qquad (4\text{-}7)$$

We can also rely on the *normalized likelihood* weights:

$$w_e = \frac{\mathcal{L}(\boldsymbol{r},\boldsymbol{c}|\boldsymbol{p}_e)}{\sum_{f=1}^{E} \mathcal{L}(\boldsymbol{r},\boldsymbol{c}|\boldsymbol{p}_f)} \quad e=1,\ldots,E. \tag{4-8}$$

for good starting values for the optimization. These weights are easy to obtain and can be used for feedback to the experts as to their predictive ability in relation to one another.

4.3 Determining the Decision Maker's Confidence in the Consensus Estimate

The decision maker's distribution, posterior to the expert opinion, is given by (4-3) with $\alpha_i = \beta \, \Sigma_{e=1}^{E} w_e n_{ie}$, $i=1,\ldots,m$. Given the restrictions on w_e, the α_0 is given by βn where n is the total subjective number of failures for the component. As with α_0 in (4-3), β controls the variability in the decision maker's probability distribution estimate. It thus acts as a measure of the decision maker's belief in the consensus estimate and controls the sensitivity of the posterior distribution to the sample information. It is difficult to elicit β directly from the decision maker though intuitively we might argue that the sum of virtual observations should be no more than n (since experts share the same information set) and thus specify the bounds of $0 < \beta \leq 1$.

We offer the following two alternatives. As β controls the belief of the decision maker in the consensus distribution we can test the decision maker's sensitivity to new information. The decision maker's probability distribution for $\boldsymbol{p}_D$ is given by (4-5) and so, for a fixed value j, we easily calculate p_{jD}. After showing this value to the decision maker, we then ask how his estimate would change if n^* new failures were observed in interval j where n^* is some suitably chosen value. Suppose the decision maker specifies a new probability $\tilde{p}_{jD}$. In the light of this new failure information, the new value of α_0 is $\beta n + n^*$ and from (4-4), the new estimate of P_{jA} is $p_{jD}^* = \mathrm{E}\{P_{jA} | p_1,\ldots,p_E, n^*\} = [\beta n/\beta n + n^*] p_{jD} + [n^*/\beta n + n^*]$. Set $\tilde{p}_{jD} = p_{jD}^*$; we solve for $\beta = n^*(1-\tilde{p}_{jD})/n(\tilde{p}_{jD} - p_{jD})$. To check the consistency of the decision maker, this is done for several values of j.

Another method for determining β involves having the decision maker specify a range R_j for any P_{jA}. This range can be equated with 6 standard deviations, and β is found by setting $R_j^2 = 6^2 \cdot \mathrm{Var}\{P_{jA}\}$ and, using the second expression in (4-4), solving $\beta = 1/n[(36 p_{jD}(1-p_{jD})/R_j^2) - 1]$. This heuristic is similar to that used for estimating parameters for activity time distributions in PERT computations.

5. UPDATING EXPERT OPINION WITH FAILURE & MAINTENANCE DATA

In updating the decision maker's probability distribution with *new data* (data obtained after the decision maker's prior distribution has been established), we use Bayes theorem. Let the new data be of the form $\boldsymbol{s} = (s_1,\ldots,s_m)$ failures and $\boldsymbol{v} = (v_1,\ldots,v_m)$ maintenance actions ($v_m \equiv 0$). The likelihood of these data is given by (4-6) with $\boldsymbol{p}_D = \boldsymbol{p}_A$, $\boldsymbol{r}=\boldsymbol{s}$, $\boldsymbol{c}=\boldsymbol{v}$. The posterior distribution is then proportional to the product of this likelihood and the prior distribution given by (4-3) with $\alpha_i = \beta \Sigma_{e=1}^{E} w_e n_{ie}$:

$$\prod_{i=1}^{m} p_{iA}^{\alpha_i + s_i - 1} \left(\sum_{j=i+1}^{m} p_{jA} \right)^{v_i}. \tag{5-1}$$

Distribution (5-1) is a *Generalized Dirichlet distribution* [24]. The Bayes estimate (posterior mean) of P_{iA}, is:

$$\mathrm{E}\{P_{iA} | p_1,\ldots,p_E, \boldsymbol{s}, \boldsymbol{v}\} = (s_i + \alpha_i) \cdot \frac{\prod_{h=1}^{i-1} \left[\sum_{z=h}^{m-1} (\alpha_{z+1} + s_{z+1} + v_z) \right]}{\prod_{h=1}^{i} \left[\sum_{z=h}^{m} (v_z + s_z + \alpha_z) \right]}, \quad i=1,\ldots,m. \tag{5-2}$$

Eq (5-2) provides a point estimate for the discretized probability distribution which represents the decision maker's updated probability distribution in light of the new data. Subsequent posterior point estimates are made by updating $\boldsymbol{s}$ and $\boldsymbol{v}$ in (5-2). Variance and covariance expressions for the P_{iA} can also be obtained. The variances can be used for establishing bounds for the distribution and the covariance provides some information on the smoothness of the posterior distribution.

6. EXAMPLE

This fictitious example illustrates our approach. Information on a particular component is elicited from four experts. Each expert provides a histogram based on a (subjective) sample of $n=100$ components over five intervals (0,2], (2,4],..., $(8,\infty)$ where the unit of time is in years. The values of n_{ie} obtained from the experts' histograms are presented in table 1. For determining the expert weights we have the following failure & maintenance data for the component: 2.00^+, 1.92, 4.00^+, 4.00^+ 6.00^+, 6.00^+, 7.69, 6.00^+, 8.00^+, and 8.00^+, where "+" indicates a censored observation as a result of a maintenance operation. Thus the existing data are $\boldsymbol{r} = (1,0,0,1,0)$ and $\boldsymbol{c} = (1,2,3,2,0)$.

TABLE 1
Elicited Values for Expert Histograms

Expert	n_{1e}	n_{2e}	n_{3e}	n_{4e}	n_{5e}
1	2	2	4	8	84
2	4	4	12	16	64
3	3	4	5	6	82
4	1	4	6	13	76

From (4-8) we derive the expert scores (0.224, 0.426, 0.200, 0.150), which can be used for feedback to the experts and as starting values in the optimization procedure for determining the experts' weights. The weights obtained from (4-7) are (0.187, 0.810, 0.001, 0.002). We thus obtain the decision maker's probability distribution using (4-5) as p_D= (0.036, 0.036, 0.105, 0.145, 0.678). Using the second heuristic approach for determining β we obtain from the decision maker the range (0,0.19) for P_{1A}, ie, $R_1 = 0.19$ and solve for $\beta = 0.3361$. Thus the total number of virtual observation for the experts as a group is approximately 34. This defines the decision maker's distribution for P_A prior to observing new data. Once new data are observed the distribution can be updated in a straightforward manner using (5-2). For example, if in the future we observe the failure and maintenance data, 4.00^+, 6.00^+, 6.00^+, 7.25, and 8.00^+, or $s = (0,0,0,1,0)$ and $v = (0,1,2,1,0)$, our revised estimate of P_A is, using (5-2), $p_D^* =$ (0.032,0.032,0.094,0.167,0.675).

ACKNOWLEDGMENT

We are pleased to acknowledge the contributions via discussions and computer programs for numerical work of Mr. Rene van Dorp.

REFERENCES

[1] R. Dekker, "Use of expert judgment for maintenance optimization", *Expert Judgment in Risk and Reliability Analysis: Experience and Perspective*, report from ESRRDA Project Group on Expert Judgment, 1989 Oct.

[2] R. M. Cooke, *Experts in Uncertainty: Expert Opinion and Subjective Probability in Science*, 1992; Oxford University Press.

[3] A. Mosleh, V. M. Bier, G. Apostolakis, "A critique of current practice for the use of expert opinions in probabilistic risk assessment", *Reliability Engineering and System Safety*, vol 20, 1988, pp 63–85.

[4] N. D. Singpurwalla, "An interactive P.C.-based procedure for reliability assessment incorporating expert opinion and survival data", *J. American Statistical Assoc*, vol 83, 1988, pp 44–51.

[5] C. S. Spetzler, C. S. Staël von Holstein, "Probability encoding in decision analysis", *Management Science*, vol 22, 1975, pp 340-358.

[6] T. S. Wallsten, D. V. Budescu, "Encoding subjective probabilities: A psychological and psychometric review", *Management Science*, vol 29, 1983, pp 151–173.

[7] H. Ibrekk, M. Granger Morgan, "Graphical communication of uncertain quantities to nontechnical people", *Risk Analysis*, vol 7, 1987, pp 519–529.

[8] I. Kabus, "You can bank on uncertainty", *Harvard Business Review*, vol 56, 1976, pp 95–105.

[9] C. Genest, J. V. Zidek, "Combining probability distributions: A critique and an annotated bibliography", *Statistical Science*, vol 1, 1986, pp 114–148.

[10] S. French, "Group consensus probability distributions: a critical survey", *Bayesian Statistics 2*, J. M. Bernardo, M. H. DeGroot, D. V. Lindley, A. F. M. Smith (eds), 1985, pp 183–202; Elsevier Science Publishers B.V.

[11] R. L. Winkler, "The consensus of subjective probability distributions", *Management Science*, vol 15, 1968, pp B61-B75.

[12] D. V. Lindley, "Reconciliation of discrete probability distributions", *Bayesian Statistics 2*, J. M. Bernardo, M. H. DeGroot, D. V. Lindley, A. F. M. Smith (eds), 1985, pp 375-390; Elsevier Science Publishers.

[13] P. A. Morris, "An axiomatic approach to expert resolution", *Management Science*, vol 29, 1983, pp 24–32.

[14] M. Mendel, T. B. Sheridan, "Filtering information from human experts", 1988; Man-Machine Systems Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology.

[15] R. L. Winkler, "Discussion on expert resolution", *Management Science*, vol 32, 1986, pp 298-303.

[16] P. Kempthorne, M. Mendel, *Adjusting Experts' Probabilistic Forecasts*, 1987; MIT Press.

[17] M. H. DeGroot, *Optimal Statistical Decisions*, 1970; McGraw-Hill Book Company.

[18] M. Stone, "The opinion pool", *Annals Mathematical Statistics*, vol 32, 1961, pp 1339-1342.

[19] K. J. McConway, "Marginalization and linear opinion pools", *J. American Statistical Association*, vol 76, 1981, pp 410-414.

[20] M. H. DeGroot, "Reaching a consensus", *J. American Statistical Association*, vol 69, 1974, pp 118-121.

[21] H. V. Roberts, "Probabilistic prediction", *J. American Statistical Association*, vol 60, 1965, pp 50–62.

[22] R. M. Cooke, M. Mendel, W. Thijs, "Calibration and information in expert resolution: A classical approach", *Automatica*, vol 24, 1988, pp 87–94.

[23] R. L. Winkler, "Scoring rules and the evaluation of probability assessors", *J. American Statistical Association*, vol 64, 1969, pp 1073–1078.

[24] R. H. Lochner, "A generalized Dirichlet distribution in Bayesian life testing", *J. Royal Statistical Society, Series B*, vol 37, 1975, pp 103–113.

AUTHORS

Ir. J. M. van Noortwijk; Faculty of Technical Mathematics & Informatics; Delft University of Technology; POBox 356; 2600 AJ Delft, The NETHERLANDS.

Jan M. van Noortwijk obtained his Masters in Applied Mathematics from the Delft University of Technology in 1989. He is working on a PhD in Applied Mathematics at Delft University of Technology. His research interests include maintenance optimization, reliability theory, and application of Bayes inference.

Dr. R. Dekker; Economics Institute; Erasmus University; POBox 1738; 3000 DR Rotterdam, The NETHERLANDS.

Rommert Dekker (A'89) is a professor of Operations Research in the Economics Institute of Erasmus University. He obtained his MA & PhD in Mathematics from the University of Leiden in 1980 & 1985. His research interests include maintenance optimization, reliability theory, and inventory theory.

Dr. R. M. Cooke; Faculty of Technical Mathematics & Informatics; Delft University of Technology; POBox 356; 2600 AJ Delft, The NETHERLANDS.

Roger M. Cooke is an Associate Professor of Operations Research and Decision Theory in the Department of Technical Mathematics and Informatics at the Delft University of Technology. He obtained his PhD in Philosophy from Yale University in 1975. His research interests include risk analysis and the resolution of expert opinion.

Dr. T. A. Mazzuchi; Department of Operations Research, School of Engineering and Applied Science; The George Washington University; Washington, DC 20052 USA.

Thomas A. Mazzuchi (A'86) is an Associate Professor of Operations Research and Engineering Management and holds a joint appointment in the Department of Operations Research and the Department of Engineering Management at The George Washington University. He obtained his MS and DSc in Operations Research from The George Washington University in 1979 & 1982, respectively. His research interests include the application of expert opinion and Bayes inference in in reliability and quality control.

Manuscript TR89-150 received 1989 September 25; revised 1991 April 20.

IEEE Log Number 06993

◄TR►