MANAGEMENT SCIENCE
Vol. 33, No. 3, March 1987
Printed in U.S.A.

COMBINING OVERLAPPING INFORMATION*

ROBERT T. CLEMEN
College of Business Administration, University of Oregon, Eugene, Oregon 97403

A decision maker trying to learn about an uncertain quantity may obtain divergent information from a number of sources (e.g., experts). In this paper we study the decision maker's problem of aggregating this information to form his posterior distribution when he believes that the experts' opinions are dependent due to shared information. Examples in both normal and Bernoulli frameworks lead to some surprising general insights regarding the impact of the dependence on the posterior distribution. The study has implications for practitioners who face real-world information-aggregation problems and for groups of experts who seek a consensus.
(BAYESIAN COMBINATION OF INFORMATION; INFORMATION AGGREGATION; CONSENSUS; DEPENDENCE)

1. Introduction

Suppose a decision maker (DM) is interested in the outcome of some uncertain event (θ). If DM acquires information about θ from a number of sources, he then faces the thorny problem of aggregating the information. In particular, if DM judges the information from the sources (experts for convenience) to be dependent, his aggregation problem can become quite complicated. In this paper we consider the problem of combining opinions from experts whose opinions are dependent because their data sets overlap.

The general problem of DM aggregating diverse information from a variety of sources has been solved, at least in principle. Morris (1974, 1977) outlines the appropriate Bayesian procedure whereby DM treats the expert opinions as data and assesses a likelihood function for these data. Application of Bayes' rule results in DM's posterior distribution of θ. However, the assessment of the likelihood function may be difficult if the information is dependent. Bayesian models for combining dependent information can be found in Winkler (1981), French (1980, 1981), Lindley (1983, 1985), Clemen (1984) and Chang (1985).

While it is clear that many factors may influence the dependence among experts' opinions, such as similar experience, training or approaches to a problem, shared information is often invoked as a primary reason for dependence. For example, Lindley (1985) states that "the most important source of correlation is the knowledge held in common [by experts]. . . ." Schervish (1986) uses shared information as a motivation for dependence among experts. Winkler (1981) develops a normal distribution example in which shared information leads to dependence among experts' opinions. In an earlier paper Zeckhauser (1971) analyzes a similar situation.

We analyze a series of examples that show how the assessment of the overlap among experts' data can affect the amount of information that is available to DM and hence his posterior distribution. The examples demonstrate some surprising insights regarding dependence resulting from shared information. Since we focus on information overlap, the examples abstract away many realistic elements of the problem; we assume the experts form their opinions in well-specified ways, that they observe random variables that are independent and identically distributed (iid), and that DM has substantial knowledge about the experts as well as the distribution of the random variables. The experts are modeled essentially as data collectors and summarizers who have access to

* Accepted by Robert L. Winkler; received March 6, 1986.

0025-1909/87/3303/0373$01.25

information not directly available to DM. Even with such simplifications, these special cases provide important insights pertinent to the development of information-aggregation models.

The next section formalizes the general approach that we take and establishes the necessary notation. §§3 and 4 analyze parallel examples of aggregating expert opinions based on shared observations from normal and Bernoulli processes, respectively. In §5 we discuss more generally the insights developed in the previous two sections. §6 concludes with a discussion of some implications for research in the area of consensus theory and modeling and for practitioners who must regularly aggregate information.

2. The General Model, Assumptions and Notation

Denote the uncertain event of interest by θ, and let $X = \{x_1, \ldots, x_n\}$ be a set of outcomes from n independent and identically distributed random variables related to θ. Expert i (E_i) observes a subset of X (X_i, containing n_i of the n outcomes in X), forms his opinion $g_i(\theta|X_i)$, and reveals g_i to DM. While g_i could be a probability distribution, in this paper it will be simply a sufficient statistic for X_i. To avoid complications we assume that E_i has a noninformative natural conjugate distribution for θ prior to observing X_i.

If DM now obtains information from E_i and E_j and wishes to use this information to form his posterior distribution for θ, then his posterior will depend on the extent to which X_i and X_j overlap. Furthermore, if k experts are consulted, DM must consider all 2-way, . . . , k-way intersections of $X_1, \ldots, X_k$. To specify the structure of the overlap among the experts' information, let $X_{ij} = X_i \cap X_j$ and in general $X_{i\ldots l} = X_i \cap \cdots \cap X_l$. Also, let $n_{i\ldots l}$ be the number of elements (outcomes) in $X_{i\ldots l}$. We will assume that DM knows both the amount of information available to each expert as well as how the experts share information. Thus, DM knows $n_1, \ldots, n_k, n_{12}, \ldots, n_{1\ldots k}$. Finally, DM uses this knowledge to determine the joint likelihood for the experts' opinions and forms his posterior distribution by applying Bayes' rule. Again for convenience we assume that DM has a noninformative distribution of θ prior to hearing the experts' information. If this is not the case, then DM may find it convenient to model his information as though it comes from a k + 1st expert observing X_{k+1}. He may then proceed with the analysis, augmenting the joint likelihood function to account for the additional information. If DM's priors cannot be modeled in this way, re-analyzing the problem with a proper, perhaps not natural-conjugate, prior distribution may be required.

3. Combining Overlapping Information: A Normal Model

Assume that each x_i is the outcome of a normal random variable with mean θ and known variance $\sigma^2 = 1$. E_i estimates θ with $\bar{x}_i$, the arithmetic average of the elements of X_i and reveals $\bar{x}_i$ to DM, who then has to find his posterior distribution $f(\theta|\bar{x}_1, \ldots, \bar{x}_k)$. Let $\bar{X} = (\bar{x}_1, \ldots, \bar{x}_k)^t$, where t denotes transposition. Conditional on θ, $\bar{X}$ can be shown to be normal with mean vector $(\theta, \ldots, \theta)^t$ and covariance matrix Σ, where Σ's ith diagonal element is $1/n_i$ and i, jth off-diagonal element is n_{ij}/n_in_j. Following Winkler (1981), DM's posterior distribution $f(\theta|\bar{X})$ is normal with parameters μ^* and σ^{*2} where

$$\mu^* = \sigma^{*2} e^t \Sigma^{-1} \bar{X}, \tag{1a}$$

$$\sigma^{*2} = 1/e^t \Sigma^{-1} e, \quad \text{and} \tag{1b}$$

$$e = (1, \ldots, 1)^t.$$

EXAMPLE 1N. DIFFERENT OVERLAP PATTERNS IN THE NORMAL MODEL. We will consider two different cases. First, suppose that $n = 3$ and $k = 3$, with $X_1 = \{x_1, x_2\}$, $X_2 = \{x_1, x_3\}$ and $X_3 = \{x_2, x_3\}$. For each expert, half of his information is shared with a second expert and the other half is shared with the third expert. In this case we can show that the correlation between any two experts is 0.5, and $\sigma^{*2} = \frac{1}{3}$. For the second case, suppose that $n = 4$ and $k = 3$, with $X_1 = \{x_1, x_4\}$, $X_2 = \{x_2, x_4\}$ and $X_3 = \{x_3, x_4\}$. Now each expert observes some information privately, but there is some information available to all experts (x_4). Performing the calculations, we find that the experts have the same correlation and that σ^{*2} again is $\frac{1}{3}$.

In these two cases the data overlap in substantially different ways. Further, it would appear in the second case that there is more independent information available with the addition of x_4. Finally, as soon as DM hears from any two experts in the first case, the remaining expert would appear to have completely redundant information. How is it, then, that σ^{*2} is the same in each case?

The apparent redundancy in the first case is explained by noting that DM can in fact deduce the values of x_1, x_2 and x_3 by solving the system of equations

$$x_1 + x_2 = 2\bar{x}_1, \qquad x_1 + x_3 = 2\bar{x}_2, \qquad \text{and} \qquad x_2 + x_3 = 2\bar{x}_3.$$

Knowing how x_1, x_2 and x_3 are distributed, DM can use them directly to find his posterior distribution with $\sigma^{*2} = \frac{1}{3}$. Thus, he is able in the first case to extract every bit of information available in X.

In the second case, if he could extract all of the information available in X, his posterior variance would be $\frac{1}{4}$. However, the fact that x_4 is observed in common by all three experts means that its presence has a confounding effect through the introduction of dependence among the $\bar{x}_i$'s. The confounding effect exactly offsets the additional information in the system. Furthermore, it is possible to show that if the variance of x_4 is greater than 1 (it contains less information than x_1, x_2 or x_3) then σ^{*2} is greater than $\frac{1}{3}$, the value for σ^{*2} if x_4 were absent altogether. In other words, it is possible to add information to the system in such a way that the overall effect is to reduce the information available to DM.

EXAMPLE 2N. IGNORING AN EXPERT IN THE NORMAL MODEL. Suppose that E_i's information completely subsumes E_j's ($X_j \subset X_i$). It seems reasonable that DM would use only the information from E_i. Unfortunately, this is only true under certain circumstances, as the following examples demonstrate. Suppose that $k = 2$ and that $X_2 \subset X_1$. Then it is straightforward to show that $\mu^* = \bar{x}_1$ and $\sigma^{*2} = 1/n_1$, which is exactly $f_1(\theta|X_1)$. E_2 plays no part whatsoever in DM's posterior beliefs and thus can be ignored. Suppose also that $X_i = X_j$. In this case $\bar{x}_i = \bar{x}_j$ and so one of the two experts can obviously be ignored.

However, a seemingly extraneous expert may actually be of some use if there are three or more experts. Suppose that $k = 3$ and that $X_1 = \{x_1, x_2, x_4\}$, $X_2 = \{x_2, x_4\}$ and $X_3 = \{x_3, x_4\}$. Again E_2's information is completely subsumed by E_1, but E_3 also shares some of E_2's information. DM's posterior parameters are $\mu^* = (9/11)\bar{x}_1 - (2/11)\bar{x}_2 + (4/11)\bar{x}_3$, and $\sigma^{*2} = 3/11$. If DM chose to ignore E_2, his posterior parameters would be $\mu^* = (2/3)\bar{x}_1 + (1/3)\bar{x}_3$, and $\sigma^{*2} = \frac{5}{18} > \frac{3}{11}$. It is possible, of course, that the values for x_1, x_2 and x_3 could be such that μ^* could be the same in each case. However, the fact that σ^{*2} is less when E_2 is considered reflects the fact that E_2 indeed adds information. The interpretation is that DM is able to use $\bar{x}_2$ to learn something about the private information that the other experts have observed.

An immediate question is whether these information effects are due to the characteristics of the normal distribution. The Bernoulli examples in the next section and the discussion in §5 show that similar phenomena occur generally.

4. Combining Overlapping Information: A Bernoulli Model

Suppose that each x_i is the outcome of a Bernoulli process with proportion p, the parameter of interest. E_i, observing s_i successes in n_i trials, computes $p_i = s_i/n_i$ and reports this to DM as his estimate of p. DM then has to find $f(p|p_1, \ldots, p_k)$. Knowing $n_1, \ldots, n_k$, DM is able to calculate $S = (s_1, \ldots, s_k)$ but, due to the information overlap, may not know s, the total number of successes in the n trials. In the appendix we show that DM's posterior distribution of p (when he has an improper diffuse prior) is given by a mixture of beta distributions:

$$f(p|S) = \sum_{q=0}^{n} v_q^* f_\beta(p|q, n), \tag{2}$$

where v_q^* can be interpreted as the posterior probability, given the experts' reports, that $s = q$, and $f_\beta(p|q, n)$ is a beta distribution with parameters q and n.

EXAMPLE 1B. DIFFERENT OVERLAP PATTERNS IN THE BERNOULLI MODEL. Suppose that $n = 6$, $k = 3$, and $X_1 = \{x_1, x_2, x_3, x_4\}$, $X_2 = \{x_1, x_2, x_5, x_6\}$ and $X_3 = \{x_3, x_4, x_5, x_6\}$. As before, each expert shares half of his information with a second expert and the other half with a third expert. Let s_{ij} be the number of successes observed in common by E_i and E_j. Now DM can solve the equations

$$s_{12} + s_{13} = s_1, \qquad s_{12} + s_{23} = s_2, \qquad \text{and} \qquad s_{13} + s_{23} = s_3$$

for the values of s_{12}, s_{13} and s_{23}, add them to obtain s, and finally apply Bayes' rule to obtain his posterior distribution $f_\beta(p|s, n)$. For example, if $S = (2, 2, 2)$, DM's posterior distribution is $f_\beta(p|s = 3, n = 6)$.

For the second case, let $n = 8$, $k = 3$, $X_1 = \{x_1, x_2, x_7, x_8\}$, $X_2 = \{x_3, x_4, x_7, x_8\}$ and $X_3 = \{x_5, x_6, x_7, x_8\}$. Again we have $n_1 = n_2 = n_3$, but now each expert observes private data as well as a core of data available to all experts. If $S = (2, 2, 2)$, DM's posterior distribution can be determined by applying (2):

$$f(p|S) = 0.147 f_\beta(p|2, 8) + 0.706 f_\beta(p|4, 8) + 0.147 f_\beta(p|6, 8).$$

Figure 1 shows the posterior distributions for these two cases. Note also that if we were to eliminate x_7 and x_8 from the second case, the posterior distribution would be the same as in the first case. The data x_7 and x_8 introduce dependence among the experts and produce a confounding effect that more than offsets the additional information, resulting in the mixed distribution being more spread out than the single distribution.

EXAMPLE 2B. IGNORING AN EXPERT IN THE BERNOULLI MODEL. Again we consider the problem of whether an expert whose information is subsumed by another can be ignored. Suppose that $k = 2$. The two experts see n_{12} trials in common, but E_1 observes n_1 trials privately. Since $n = n_1$, then $s = s_1$, and DM's posterior distribution is $f_\beta(p|s_1, n_1)$. Once again it is obvious that E_2 plays no part in DM's posterior beliefs.

However, as in the normal case, this does not hold generally. Suppose $n = 8$, $k = 3$, $X_1 = \{x_1, \ldots, x_6\}$, $X_2 = \{x_1, \ldots, x_4\}$ and $X_3 = \{x_1, x_2, x_7, x_8\}$, so that all of E_2's data are observed by E_1, but now E_3 also observes some of E_2's data. For this example, suppose that $s_1 = 3$ and $s_3 = 2$. If DM ignores E_2, his posterior distribution is

$$f(p|s_1 = 3, s_3 = 2) = 0.154 f_\beta(p|3, 8) + 0.692 f_\beta(p|4, 8) + 0.154 f_\beta(p|5, 8).$$

Given that $s_1 = 3$ and $s_3 = 2$, s_2 can be 1, 2 or 3. If E_2 is not ignored and $s_2 = 1$, then DM's posterior distribution is

$$f(p|s_1 = 2, s_2 = 1, s_3 = 2) = 0.6 f_\beta(p|4, 8) + 0.4 f_\beta(p|5, 8).$$

Similarly for $s_2 = 2$ we have

$$f(p|s_1 = 3, s_2 = 2, s_3 = 2) = 0.125 f_\beta(p|3, 8) + 0.750 f_\beta(p|4, 8) + 0.125 f_\beta(p|5, 8),$$

and for $s_2 = 3$

$$f(p|s_1 = 3, s_2 = 3, s_3 = 2) = 0.4 f_\beta(p|3, 8) + 0.6 f_\beta(p|4, 8).$$

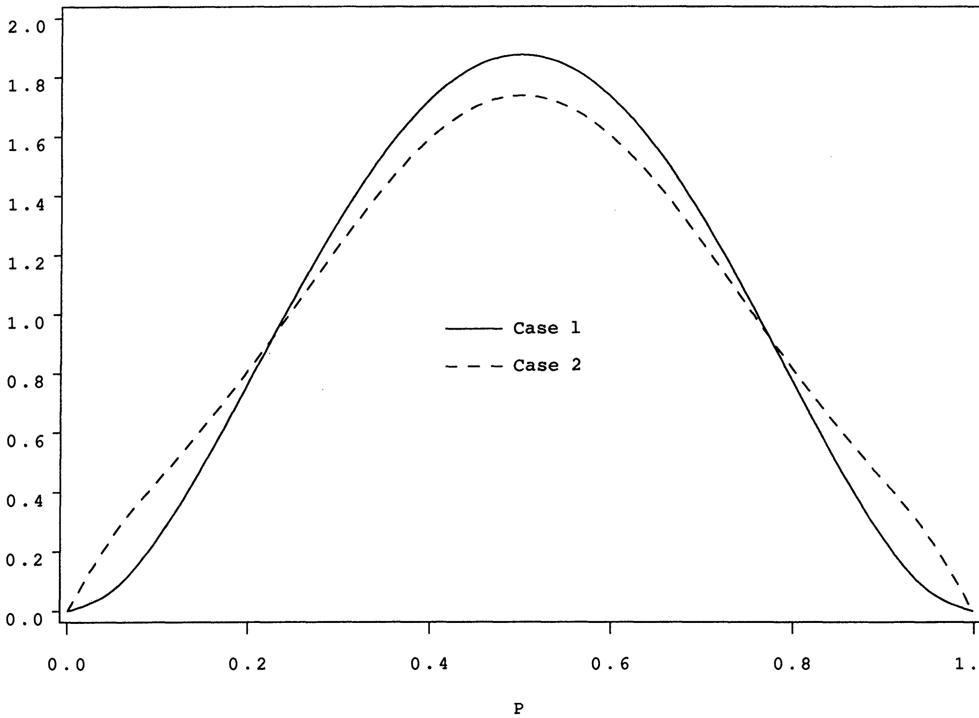

FIGURE 1. Two possible posterior distributions for p when there are three experts, each observing four trials and reporting $s_1 = s_2 = s_3 = 2$. The two cases are described in the text. In Case 1, the data are shared in a way that allows DM to deduce s, the total number of successes in the data set. In Case 2, the addition of a commonly observed set of observations keeps DM from being able to calculate s.

It is also interesting to note the relationships among the posterior expectations of p, $E(p|s_1, s_2, s_3)$. In all cases $p_1 = p_3 = 0.5$. However, $s_2 = 1$ implies that $p_2 = 0.25$, while $s_2 = 2$ implies $p_2 = 0.50$ and $s_2 = 3$ implies $p_2 = 0.75$. Thus, calculating the posterior expectations for the three distributions above when E_2 is consulted gives:

$$E(p|p_1 = 0.50, p_2 = 0.25, p_3 = 0.50) = 0.55,$$

$$E(p|p_1 = 0.50, p_2 = 0.50, p_3 = 0.50) = 0.50, \qquad \text{and}$$

$$E(p|p_1 = 0.50, p_2 = 0.75, p_3 = 0.50) = 0.45.$$

These results are similar to those in the normal example in which E_2's information entered the DM's posterior mean with negative weight. In this case the reason for this effect is apparent. A low (high) value for s_2 means that there are few (many) successes seen in common by the experts; hence, E_1 and E_3 must have seen relatively many (few) successes in their private data.

5. Some General Remarks

Examples 1N and 1B above demonstrate that information can be added to an information system in such a way that less information is available to DM. In general, represent the underlying data and the experts' opinions as vectors $X = (x_1, \ldots, x_n)$ and $G = (g_1, \ldots, g_k)$, respectively. It is clear that G is simply a function of X, say $\gamma(X)$. If γ is 1-to-1, then G and X are equivalent in terms of the information that is available to DM. However, if γ is not 1-to-1, as may occur in the case of information overlap, then G is less informative than X. Alternatively, X is sufficient for G in the sense of Blackwell

(1953). Furthermore, it is also obvious that different transformations of X, say $\gamma_1(X) = G_1$ and $\gamma_2(X) = G_2$, may have different information content for DM; that is, G_1 may be less informative than G_2, which in turn may be less informative than X.

The examples have demonstrated that additional information may actually be of negative value to DM. At first glance this appears to be somewhat paradoxical; a standard result in statistical decision theory (e.g., LaValle 1968) states that information always has nonnegative value. However, this result applies to information that is available to DM. In our examples we have shown that it is possible to change the information system in such a way that less information is available to DM; that is, while we add information to the system, we modify the system at the same time so that DM has more trouble "sorting out" what information there is in the experts' opinions.

Examples 2N and 2B examine the general area of sufficiency in the context of the experts themselves rather than the underlying information. When there were two experts, one could be ignored if his data set were subsumed by another expert. This is clearly true in general when the experts report sufficient statistics for their data. However, consider three expert opinions g_1, g_2, g_3. It is obvious that obtaining information from all three experts is always sufficient for information from any two experts. Are there conditions, though, where information from two of the experts is sufficient for all three? There are at least two possibilities: (1) two of the experts observe the same information and process it the same way, or (2) one expert's data set is subsumed by the second expert, and the third expert's opinion is conditionally independent of the first expert's opinion.

For the first possibility, it is obvious why one of the experts can be ignored. In the second case, suppose that $X_1 \subset X_2$. The sufficient statistic reported by E_2 is sufficient for the information in X_1. Thus

$$f(\theta|g_1, g_2) = f(\theta|g_2) \qquad \text{and} \qquad f(\theta|g_1, g_2, g_3) \propto f(g_3|\theta, g_1, g_2)f(\theta|g_2).$$

If g_3 is independent of g_1 (conditional on θ and g_2), then $f(g_3|\theta, g_1, g_2) = f(g_3|\theta, g_2)$. This yields

$$f(\theta|g_1, g_2, g_3) \propto f(g_3|\theta, g_2)f(\theta|g_2) \propto f(\theta|g_2, g_3),$$

which demonstrates that E_1 can indeed be ignored by DM. The conditional independence condition is met, for example, if the x_i are iid and there is no overlap between X_3 and X_2. (In this case, g_3 is conditionally independent of both g_1 and g_2.)

These arguments can apply to situations considerably more general than the simple examples analyzed in §§3 and 4. The x_i need not be iid, we can readily extend the models to more than three experts, and DM may have more vague information about the overall information system (process parameters, overlap structure, etc.). Such enhancements complicate the Bayesian analysis leading to DM's posterior, but the analysis is still relatively straightforward. For details, see Clemen (1984).

6. Conclusion

The examples we have analyzed lead to some rather remarkable insights. A group of experts may observe a new bit of information in common that actually reduces the overall information available to the DM; in general there is a tradeoff between the confounding effect of a new shared observation and the additional information that it brings to the system. Not so surprising, perhaps, is the notion that different kinds of information overlap lead to different ways of aggregating the experts' opinions; it is obvious on reflection that changing the nature of the dependence among a set of experts will lead to a new aggregation formula.

While the examples and analysis have been rather abstract, there are some obvious implications for individuals who face real-world aggregation problems. One implication is that overlapping information may not be terribly useful; one may wish to assemble a set of experts whose information overlaps less, on the grounds that there will be more information available from the group whose information is less interdependent. [See Clemen and Winkler 1985 for a related analysis that leads to the same conclusion.] In the case of aggregating economic forecasts, say, one might like to select a group of forecasters who exhibit as little dependence as possible. Such a group would be a heterogeneous collection of forecasting approaches, theoretical persuasions, and hence different views of available data. For example, Robert J. Eggert publishes Blue Chip Economic Indicators, a set of consensus forecasts, each composed of over 40 different individual forecasts. In an interview with the author, Mr. Eggert has indicated that he tries to identify "independent thinkers" to include in his group; this reflects a practitioner's informal approach to the general problem of assembling a "less dependent" group of experts.

An abundance of research has focused on how a group of experts should combine their opinions to arrive at a single consolidated opinion without the aid of an outside decision maker. French (1985) and Genest and Zidek (1986) provide critical reviews. It is important to realize that an assessment of the dependence among the expert opinions may be an important aspect of the aggregation. In the group consensus problem, though, it is not clear how such assessment should be performed or by whom. Effectively, though, we must require the experts to agree on the nature of the stochastic process that jointly generates their opinions and θ.

Following the discussion in §5, a consolidated single opinion from the group will likely be less informative for DM than the set of individual opinions. The only case in which this is not true is when DM's posterior would be the same whether he receives the consensus opinion or the individual opinions, an unlikely event requiring both the experts and DM to assess the dependence among the experts' opinions in equivalent ways. Furthermore, different decision makers may assess the dependence among the experts in different ways, and hence may combine the experts' opinions differently. Another way to look at the problem is that experts who seek consensus must report a sufficient statistic for their opinions. Without widespread agreement about the stochastic process that jointly determines their opinions and θ, this task makes no sense. While there are many motivations for forming a consensus, the results presented here show clearly why a group of experts should report their individual opinions as well as any consensus view. From an information perspective, reporting only the consensus limits the information available to the eventual user of the information.[1]

[1] This research was supported in part by the National Science Foundation under Grant IST 8600788. The author gratefully acknowledges comments on early versions of this paper by Robert L. Winkler, as well as those provided by an Associate Editor and anonymous referees. As usual, the author must take the blame for any remaining errors.

Appendix

To derive the DM's posterior distribution for the Bernoulli model, let $N^* = (n_1^*, \ldots, n_k^*, n_{12}^*, \ldots, n_{1\cdots k}^*)$, where $n_{i\cdots l}^*$ is the number of observations seen only by experts $i \cdots l$. Since DM knows the overlap structure, he knows N^*. Also, let $S^* = (s_1^*, s_2^*, \ldots, s_{123\cdots k}^*)$, where $s_{i\cdots j}^*$ is the number of successes in the $i \cdots j$th sample with sample size $n_{i\cdots j}^*$. (Thus s_i^* would be the number of successes in expert i's private observations.) Then

$$\mathcal{L}(S|p, N^*) \propto \sum_{s_1^*=0}^{n_1^*} \cdots \sum_{s_{1\cdots k}^*=0}^{n_{1\cdots k}^*} \mathcal{L}(S|p, N^*, S^*)P(S^*|p, N^*).$$

The conditional likelihood in the RHS of this equation is 1 if S is consistent with S^* (that is, if the $s_{i\cdots l}^*$'s add up correctly for each s_j), and 0 otherwise. The probability $P(S^*|p, N^*)$ is simply the joint probability of a number of non-overlapping and hence independent binomial samples. Let

$$\delta(Q) = \begin{cases} 1 & \text{if expression } Q \text{ is true,} \\ 0 & \text{otherwise,} \end{cases}$$

$$I = \{1, \ldots, k, 12, \ldots, 1k, \ldots, 123\cdots k\},$$

and $\Gamma\{R^*\}$ be a transformation that maps a vector R^* into a vector (R) of successes reported by the experts according to the way their observations overlap. Then

$$\mathcal{L}(S|p, N^*) \propto \sum_{s_1^*=0}^{n_1^*} \cdots \sum_{s_{1\ldots k}^*=0}^{n_{1\ldots k}^*} \delta(S = \Gamma\{S^*\}) \prod_{i\in I} \binom{n_i^*}{s_i^*} p^s(1-p)^{n-s}.$$

By assumption, DM's prior is $f(p) \propto p^{-1}(1-p)^{-1}$. As long as $s_j > 0$ for some j and $s_i < n_i$ for some i, the DM's posterior distribution is proper and is

$$f(p|S^*, N^*) \propto \sum_{s_1^*=0}^{n_1^*} \cdots \sum_{s_{1\ldots k}^*=0}^{n_{1\ldots k}^*} u_{S^*} f_\beta(p|s, n) \qquad \text{where}$$

$$u_{S^*} = u_{s_1^*,\ldots,s_{1\ldots k}^*} = \delta(S = \Gamma\{S^*\}) \prod_{i\in I} \binom{n_i^*}{s_i^*} \mathrm{B}(s, n), \qquad f_\beta(p|s, n) = B(s, n)^{-1} p^{s-1}(1-p)^{n-s-1},$$

and $B(s, n)$ is the beta function. If we let

$$v_q^* = v_q / \sum_{Y=0}^{n} v_Y, \qquad v_q = \sum_{s_1^*=0}^{n_1^*} \cdots \sum_{s_{1\ldots k}^*=0}^{n_{1\ldots k}^*} \delta(q = s) u_{S^*}, \qquad \text{and} \qquad q \in \{0, \ldots, n\},$$

then $f(p|S, N^*) = \sum_{q=0}^{n} v_q^* f_\beta(p|q, n)$.

References

BLACKWELL, D., "Equivalent Comparisons of Experiments," *Ann. Math. Statist.*, 24 (1953), 265–272.

CHANG, K., "Combination of Opinions: The Expert Problem and the Group Consensus Problem," Ph.D. Dissertation, University of California, Berkeley, 1985.

CLEMEN, R., "Modeling Dependent Information: A Bayesian Approach," Ph.D. Dissertation, Indiana University, 1984.

———, AND R. WINKLER, "Limits for the Precision and Value of Information from Dependent Sources," *Oper. Res.*, 33 (1985), 427–442.

FRENCH, S., "Updating of Belief in the Light of Someone Else's Opinion," *J. Roy. Statist. Soc. Ser. A*, 143 (1980), 43–48.

———, "Consensus of Opinion," *European J. Oper. Res.*, 7 (1981), 332–340.

———, "Group Consensus Probability Distributions: A Critical Survey," in *Bayesian Statistics 2*, J. Bernardo et al. (Eds.), North-Holland, Amsterdam, 1985, 183–197.

GENEST, S. AND J. ZIDEK, "Combining Probability Distributions: A Critique and an Annotated Bibliography," *Statistical Sci.*, 1 (1986), 114–148.

LAVALLE, I., "On Cash Equivalents and Information Evaluation in Decisions Under Uncertainty," *J. Amer. Statist. Assoc.*, 63 (1968), 252–290.

LINDLEY, D., "Reconciliation of Probability Distributions," *Oper. Res.*, 31 (1983), 866–880.

———, "Reconciliation of Discrete Probability Distributions," in *Bayesian Statistics 2*, J. Bernardo et al. (Eds.), North-Holland, Amsterdam, 1985, 375–390.

MORRIS, P., "Decision Analysis Expert Use," *Management Sci.*, 20 (1974), 1233–1241.

———, "Combining Expert Judgements: A Bayesian Approach," *Management Sci.*, 23 (1977), 679–693.

SCHERVISH, M., "Combining Expert Judgments," *Management Sci.*, 32 (1986), 306–312.

WINKLER, R., "Combining Probability Distributions from Dependent Information Sources," *Management Sci.*, 27 (1981), 479–488.

ZECKHAUSER, R., "Combining Overlapping Information," *J. Amer. Statist. Assoc.*, 66 (1971), 91–92.