# Combining the Opinions of Experts who Partition Events Differently

Robert F. Bordley
General Motors Research Labs & University of Michigan, Ann Arbor
525 Choice Court; Troy, Michigan 48085
robert.bordley@gm.com

October 28, 2008

**Abstract**

This paper focuses on updating a client's beliefs about an event based on information about the different probabilities which various experts assess for that event. A substantial literature solves this problem when all experts assess their probabilities over the same partitioning of the possible outcomes of an event. But different experts often think about the same problem in quite different ways. This can lead to differences in how experts prefer to partition the possible outcomes of an event. Forcing the experts to use a common partition could lead to less informative probability assessments. As a result, this paper presents a new approach for combining probability assessments from different experts which allows experts to assess their probability assessments across different partitionings.
Key words: probability elicitation, decision analysis, forecasts: combining; probability:group; incoherence
History: Received on April 4,2008; Accepted on September 25, 2008 following 2 revisions.

## 1 Introduction

Decision makers often need to combine probabilistic information from different experts. But sometimes the process of combining this information is complicated by differences in how the experts think about the events of interest. For example, suppose an automotive manufacturer has an initial probability assessment over the various reasons leading some particular customer to have a vehicle serviced. The client wishes to update that prior probability based on probabilities elicited from the following four experts:

- One expert has decades of experience in observing customer complaints and decomposes the reasons for service into such categories as: *engine won't start, windshield wiper squeaks, low-level vehicle noise, door handle fell off, flashing warning signal on instrument panel,* etc.

- A second expert has decades of experience fixing vehicle problems and decomposes the reasons for service into such categories as: *gas cap not screwed on properly, wires not properly grounded, leaves in the blower motor, misaligned pins in the ignition switch,* etc.

- A third expert has decades of experience designing vehicles and decomposes the reasons for service into such categories as: *engine system, transmission system, fuel system, braking system, steering and suspension system, heating and cooling system, electrical system,* etc.

- A fourth expert has decades of experience identifying which organization created the customer's problem and decomposes the reasons for service into such categories as: *inadequacy in product design, supplier parts not meeting design specifications, part damaged in shipment to the assembly plant, wrong parts used in assembly plant, error in part assembly, inappropriate customer use of vehicle,* etc.

This example highlights how categorizations often vary with the expert's experience and with the purposes for which the expert uses the categorization. For simplicity, suppose that each customer has their vehicle serviced for one and only one reason. Then these four different categorizations (or languages (Keisler and Keisler,1999)) can be used to define four very different partitionings of the possible reasons for a vehicle being serviced. (Thus for each expert, each vehicle's reason for being serviced is associated with one and only one element of that expert's partitioning.)

There exist well-established formulas for updating the client's beliefs given probability assessments from multiple experts (French, 1985; Genest & Zidek, 1986; Clemen & Winkler, 1999; Predd et al, 2008). These formulas presume that all experts use the same partitioning, i.e., the same categorization of possible reasons for vehicle service. But each expert's experience was based on very different categorizations of reasons for service (and thus very different partitionings). Hence to apply existing formulas for aggregating expert probabilities, we must require all experts to use a common partition (and thus a partition involving categorizations different than the categorizations the experts used during their years of experience.)

This may make it harder for experts to draw on their previous experience in assessing probabilities. (Thus an expert experienced with customer complaints might have considerable intuition about the frequency of low-level noise problems but much less intuition about the frequency of problems due to leaves in the blower motor.) Reducing the amount of experience used in assessing probabilities not only makes them less informative but also makes them less calibrated (Wallsten and Budescu, 1980) and more susceptible to partition-dependence biases (Fox & Rottenstreich, 2003; Clemen & Fox, 2005; Fischhoff, Slovic & Lichtenstein, 1978; Johnson, Hershey, Meszaros & Kunreuther, 1993; Fischhoff, Slovic & Lichtenstein, 1978; Fox & Tversky, 1995; Rottenstreich & Tversky, 1997). As a result, this paper allows each expert to use their preferred partitioning in assessing probabilities.

Allowing the experts to use their preferred partitioning is consistent with the common practice of making the expert's assessment task as easy as possible. Thus Morgan and Henrion (1990, pg.143) write that *"in the process of defining the quantity that will be the focus of the assessment, it is also important to choose a form corresponding to the nature of the expert's knowledge. The expert should not need to engage in units conversion or other mental gymnastics to answer the questions that will be posed."* In describing their own elicitation work, Morgan and Henrion (1990, pg.148) also write that *" we were careful to make it clear to each atmospheric expert that we were prepared to build a model that reflects his views and that he should specify the structure with which he felt most comfortable. Thus, for example, while all experts chose to express their views about average oxidation rates in terms of first-order rate coefficients, they were explicitly told they need not do so."*

If the experts are allowed to assess probabilities over their preferred partitions, then a new method is required to aggregate expert probabilities assessed across different partitions. This method will require the introduction of a detailed partitioning that includes all the expert partitionings as special cases. It also draws on four important ideas from the existing literature:

- We follow Clemen (1987, pg. 373) in treating experts *"as data collectors and summarizers who [have] access to information not directly available to the decision maker."*
- We use Good (1950)'s method of imaginary results to elicit not only the expert's probability assessment but additional information as well.
- We apply Dickey, Jiang and Kadane (1987)'s work on pooling observations from multiple surveys, defined across different partitions, to

construct likelihood functions for the information from the experts.

- We use the fact, commonly applied in preposterior analysis, that the decision maker's posterior probability, given some information, is uncertain before the decision maker actually receives this information. We adapt Lindley, Tversky and Brown (1979)'s treatment of a decision analyst as being uncertain about a client's subjective probability to define a probability over this uncertain posterior probability.

The next section describes how these four key ideas are used to develop a formula for combining expert probabilities assessed over different partitions. The paper closes with an illustration of the approach on an automotive example.

# 2 The Aggregation Formula

## 2.1 Experts as Information Sources

Following Clemen (1987), we treat expert $k$ as if the expert's knowledge about the occurrence of event $E$ was purely based on recording the number of instances of event $E$ in a subpopulation (not accessible by the client) in which there were $n_k$ total instances of some event occurring. If the expert observes $n_k(E)$ occurrences of event $E$, then $\sum_{E \in B_k} n_k(E) = n_k$. Note that observing $n_k(E)$ occurrences of event $E$ for each $E$ in $B_k$ involves classifying each of the $n_k$ observations using one of the categories in the partition. Hence we could think of the expert as answering a multiple-choice question about each of the $n_k$ occurrences (with each possible response to the multiple-choice question corresponding to a separate category in the partition.) This makes the information collected by the expert analogous to the information arising from $n_k$ surveys. Receiving information from experts using different partitionings then corresponds to receiving information from different surveys using a multiple-choice question in which the possible responses to the multiple-choice question are defined differently.

## 2.2 Eliciting the Expert's Information

To extract as much information from the expert as would be assessed from a comparable survey, the decision analyst must infer $n_k(E)$ for each $E \in B_k$. A conventional approach to probability assessment might elicit the expert's subjective probability, $P_k(E)$, compute $n_k(E) = P_k(E)n_k$ and then have the expert assess $n_k$. But since experts do not think of themselves as data collectors, they may not be able to meaningfully assess $n_k$. This reflects the fact that $n_k(E)$ and $n_k$ are actually hyperparameters of the expert's probability which are not observed by the expert at all. Fortunately, *it is often possible to choose reasonable values for the hyperparameters by appealing to the 'device of imaginary results'* (Schlaiffer, 1997).

Instead of assessing $n_k(E)$ by eliciting the expert's subjective probability for event $E$, our variation on Good (1950)'s *device of imaginary results* elicits the answer to the following question:

> **Suppose you learned of one new case where an event comparable to event $E$ occurred. Given this new information, what would be your assessed probability for event $E$ happening?**

Experts can easily understand this question for events analogous to successive flips of a possibly biased coin. But more effort is required to make the question comprehensible when the 'event comparable to event $E$' is an unrelated event which, nonetheless, the expert would regard as part of the same exchangeable series of events as event $E$.

We elicit expert $k$'s response to this question for each event $E \in B_k$. Thus if we were concerned with whether a possibly biased coin, if flipped, would come up heads, we would ask the subject to both assess the probability of heads (if he/she had previously observed a coinflip which yielded heads) and to assess the probability of tails (if he/shey had previously observed a coinflip which yielded tails.) Define $P_k(E|D_E)$ as the expert's answer to this question, i.e., the expert's assessed probability for event $E$ given an observation of a comparable event occurring. Since $P_k(E|D_E) = (n_k(E) + 1)/(n_k + 1)$, we can compute $n_k(E)$ using the formula

$$n_k(E)+1 = (n_k+1)P_k(E|D_E) \quad \text{where} \quad n_k+1 = \frac{|B_1| - 1}{(\sum_{E \in B_1} P_{E|D_E}) - 1} \tag{1}$$

Note that $\frac{1}{n_k+1}$ is the average amount by which the expert changes the probability assessment in response to a single observation. Since our interpretation of $n_k(E)$ in terms of observations is merely a useful fiction, we will not require that $n_k(E)$ be an integer.

### 2.3 Specifying a Reference Partition

Each expert $k$ has his or her preferred partition $B_k$. Prior to learning of the expert information, the client also has a preferred partition, $B_0$, over which the client, if requested, would prefer to assess the prior probability. We now use standard procedures (e.g., Billingsley, 1995) to specify a reference partition, $B$, as the coarsest possible partitioning which includes every event in any of the expert or client partitions. In the automotive problem, one element of partition $B$ might be

- the occurrence of low-level vehicle noise (from the first expert's categorization),
- associated with the climate control system (from the third expert's categorization),
- resulting from a failure to design the hood of the vehicle properly (from the fourth expert's categorization),
- so that leaves accumulating on the windshield fall into the blower motor compartment (from the second expert's categorization).

Appropriately combining this category with other categories gives the event *low-level vehicle noise* in the first expert's partition while combining this category with different categories gives the event *failure arising from product design,* in the fourth expert's partition.

Since the client's prior probability, *before learning of the expert information,* will be most easily defined over the client's preferred partition, $B_0$, we will focus on the, as yet unspecified, posterior probability that the client would assess *after learning of the expert information.* This posterior probability will be defined over $B$. Because the client is uncertain about the information the experts provide, the client is uncertain about how the posterior probability will differ from the prior probability. Preposterior analysis (Berger, 1985) recognizes this uncertainty in determining the kinds of information to collect. For this paper, we wish to view this uncertainty about the expert information as inducing an uncertainty about the posterior probability (consistent with the suggestions of Mosleh and Bier, 1996).

To quantify this uncertainty, we follow Lindley, Tversky and Brown (1979) in focusing on this uncertainty from the perspective of the decision analyst. Specifically, the decision analyst, prior to learning of the expert's information, will be uncertain about the numbers that would be elicited for a client's posterior probability in a (possible hypothetical) elicitation exercise conducted *after the client learned about the expert's opinions.* This uncertainty reflects both uncertainty about the expert information as well as uncertainty about any prior beliefs the client may have. Since the decision analyst's uncertainty about the numbers arising from the elicitation exercise satisfy the clairvoyant test, we can meaningfully define the decision analyst's subjective probability over the posterior probability, $p_A, A \in B$. We assume that this subjective probability is a Dirichlet distribution with parameters $n_0(A), A \in B_0$ with the decision analyst's mean value for $p_A$. In assessing these parameters, the decision analyst could potentially treat the client in the same way that experts are treated and use Good's device of imaginary observations to elicit values $n_0'(E)$ from the client for each $E \in B_0$. In some cases, the analyst might considering setting $n_0(E) = n_0'(E)$ for $E \in B_0$ and $n_0(A) = 0$ otherwise.

### 2.4 Updating Beliefs using the Expert Information

Since the information provided by the experts is treated as comparable to survey information, the problem of aggregating expert information becomes a problem of pooling information from different surveys. As a result, the likelihood of expert $k$ providing the information $n_k(E)$ for each $E \in B_k$ (given the probabilities $p_A$) can be described by the multinomial distribution:

$$\prod_{E \in B_k} (P(E))^{n_k(E)} \text{ with } P(E) = \sum_{A \in E} p_A, A \in B$$

We temporarily assume the experts are independent, i.e., that their observations are drawn from different non-overlapping subpopulations and that the analyst, based on feedback from the client, does not feel any need to adjust or weight the expert assessments. These assumptions will be eliminated in the next section. Then the likelihood of eliciting $n_j(E)$ for each $E \in B_j$ for experts $j = 1, ..., K$ given the probabilities $p_A, A \in B$, is

$$\prod_{k=1}^{K} \prod_{E \in B_k} (P(E))^{n_k(E)} \quad \text{with } P(E) = \sum_{A \in E} p_A, A \in B$$

If we define $n_k(E) = 0$ when $E \notin B_k$ and define $n(E) = \sum_{k=1}^{K} n_k(E)$, then this can be written as

$$\prod_{E \subset B} (P(E))^{n(E)} \quad \text{with } P(E) = \sum_{A \in E} p_A, A \in B \tag{2}$$

These likelihood functions are conditioned on $p_A, A \in B$.

Multiplying the analyst's prior by the likelihood function in equation (2) gives a posterior probability for $p_A$ given the expert information, which is proportional to the probability in equation(2) where $n(A)$ is replaced by $n(A) + n_0(A)$. This probability is the *extended Dirichlet distribution* which Dickey, Jiang and Kadane (1987) introduced for the analysis of missing survey data. It describes the decision analyst's revised probability of the client's posterior probability $p_A, A \in B$.

## 2.5 Adjusting for Dependency Among Experts

We now eliminate the assumption of the experts being independent and also allow the client and analyst to adjust and weight the expert assessments. To do this, we follow common practice (Gelman, Carlin, Stern and Rubin, 1995) in introducing a prior distribution over the 'hyperparameters' $n_1(E), ...n_K(E)$ for each $E \subset B_0$ and integrating equation (2) over this distribution. Even though $n_k(E)$ is non-negative, we will, for convenience, assume this prior distribution is Gaussian. (An alternative way of incorporating dependency is to use Wong's (1998) extended Dirichlet distribution). Let $\mu_k(E)$ be the mean of $n_k(E)$ and $C_{(k,E),(j,E^*)}$ be the covariance between $n_k(E)$ and $n_j(E^*)$ for all experts $j, k$ and all events $E, E^* \subset B$.

We now rewrite equation(2) as

$$\exp(\sum_{k,E\subset B} n_k(E)\ln(P(E))) \text{ with } P(E) = \sum_{A\in E} p_A, A \in B$$

We integrate this equation over the Gaussian prior to get

$$\begin{aligned}\exp(\sum_{k,E\subset B} \mu_k(E)\ln(P(E)) &+ \frac{1}{2}\sum_{k,j,E,E^*\subset B_0} \ln(P(E))C_{(k,E),(j,E^*)}\ln(P(E^*))) \\ \text{with } P_E &= \sum_{A\in E} p_A\end{aligned}$$

Define $\mu(E) = \sum_{k=1}^{K}\mu_k(E)$, $C_{E,E^*} = \sum_{k=1}^{K}\sum_{j=1}^{K} C_{(k,E),(j,E^*)}$ and $L_E = \ln(P_E)$ so that the expression can be rewritten more compactly as

$$\exp(\sum_{E\subset B}\mu(E)L_E + \sum_{E,E^*\subset B_0} L_E\frac{C_{E,E^*}}{2}L_{E^*} \text{ with } L_E = \ln[\sum_{A\in E} p_A]$$

or

$$\sum_{E\subset B}(P(E))^{\mu_E - \frac{1}{2}\sum_{E^*\subset B_0} C_{E,E^*}L_{E^*}} \text{ with } P_E = \sum_{A\in E} p_A \qquad (3)$$

Hence the informativeness of the experts decreases as the correlation among experts increases.

Existing Monte Carlo/Markov Chain techniques, which can be used (Dickey, Jiang and Kadane, 1987) to estimate the moments of equation (2) when experts are independent, can also be used to estimate the moments of the more complicated distribution (equation (3)) when experts are dependent. In this paper, the first moment corresponds to the analyst's updated estimate of the client's posterior probability.

# 3 Simple Numerical Examples

## 3.1 Background for the Example

Suppose we think of the automobile as consisting of three main subsystems:

- A Frame (including wheels, chassis, suspension and brakes)
- A Propulsion System (including engine, powertrain controls, transmission, battery and gas tank) and
- A Body (including seats, doors, windows, body panels, and the instrumental panel.)

Before the early 1960's, these subsystems were first manufactured separately and then integrated by mounting the propulsion system on the frame, lowering the body over the propulsion system and finally attaching the body to the frame.

Now consider three experts (see Figure 1) who have very different ways of thinking about the vehicle. Expert 1, who might be an engineer, thinks of vehicles as having three possible problems:

- Problems that can be addressed by fixing (or replacing) components in the frame,
- Problems that can be addressed by fixing (or replacing) components in the propulsion system, and
- Problems that can be addressed by fixing (or replacing) components in the body.

Expert 2, who might be a salesperson, thinks of vehicles as having two possible problems

- Drivability problems (often due to the frame or propulsion system) and
- Comfort, convenience and appearance problems (often due to the body.)

We simplify the problem by assuming that comfort, convenience and appearance problems can all be addressed by fixing problems in the body while drivability problems can all be addressed by fixing problems in the frame or propulsion system.

Expert 3, who might be an executive, only thinks in terms of the supplier organization responsible for fixing the problem. In the automotive industry,

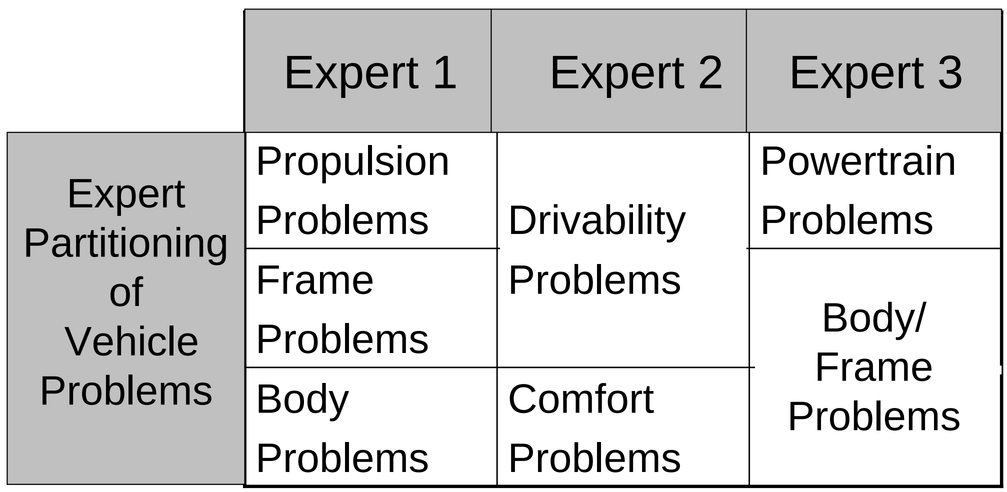


Figure 1: How 3 Different Experts Partition Vehicle Problems

it is common to have one automotive division responsible for the design and manufacture of the propulsion system and a second division responsible for the design and manufacture of body and frame. For example, Honda's Powertrain Group manufactures engines for motorcycles, snowblowers and other products totally unrelated to automobiles. This decomposition, of course, still represents an oversimplification since companies also have a general assembly division responsible for assembling the major subsystems of the vehicle.

But given this oversimplification, expert 3 will categorize problems as

- Problems originating from the Powertrain Division, and
- Problems originating from the Body-Frame Division.

We use the letters $E$, $F$ and $B$ to label the propulsion-system-based, the frame-based and the body-based problems respectively. Then expert 1's partition is $\{E, F, B\}$ while expert 2's partition is $\{(E, F), B\}$ and expert 3's partition is $\{E, (F, B)\}$ . Note that every problem occurrence (or event) in expert 2's partition (and every event in expert 3's partition) can be defined in terms of problem occurrences (or events) in expert 1's partition. On the other hand, expert 2 and expert 3 categorize events completely differently, i..e., none of the problem occurrences definable in expert 2's partition are definable in expert 3's partition.

### 3.2 Examples

We now consider two examples. In the first example, traditional approaches (which require that all experts use the same partition) are applicable; in the second example, traditional approaches are not applicable. Our proposed approach offers advantages over the traditional approach in both cases.

In the first example, the decision maker is only advised by expert 1 and expert 2. Expert 1's information consists of the values $n_1(E), n_1(F)$ and $n_1(B)$ while expert 2's information consists of the values $n_2(E, F)$, $n_2(F)$. (See Figure 2). If $n_2(E, F)$ is an integer, then, as Appendix I shows, this information can be used to compute an updated probability for the problem being a propulsion system problem (labelled $E$) as well as an updated probability for the problem being a frame problem (labelled $F$). We then use $P_B = 1 - P_E - P_F$ to compute the probability of a body problem (labelled $B$). As a result, we find

$$\begin{aligned} P_E &= \frac{1 + n_1(E)}{3 + n_1(E) + n_1(F) + n_1(B) + n_2(B) + n_2(E, F)} [1 + \frac{n_2(E, F)}{2 + n_1(E) + n_2(F)}] \\ P_F &= \frac{1 + n_1(F)}{3 + n_1(E) + n_1(F) + n_1(B) + n_2(B) + n_2(E, F)} [1 + \frac{n_2(E, F)}{2 + n_1(E) + n_1(F)}] \\ P_B &= \frac{1 + n_1(B) + n_2(B)}{3 + n_1(E) + n_1(F) + n_1(B) + n_2(B) + n_2(E, F)} \end{aligned}$$

The traditional approach would have required expert 1 to use the same partition as expert 2. This would have led to estimates of $P_E + P_F$ and $P_B$ which — depending upon the particular aggregation procedure used — could be consistent with the estimates of $P_E + P_F$ and $P_B$ given by our approach. But in addition to assessing $P_B$ and $P_E + P_F$, our approach leverages the added information provided by expert 1 to assess both $P_E$ and $P_F$.

In the second case, the decision maker is only advised by expert 2 and expert 3. Thus expert 2 provides information $n_2(E, F)$ and $n_2(B)$ while ex-

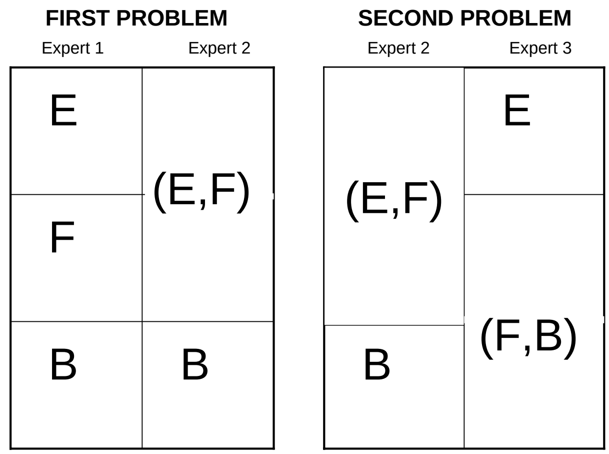


Figure 2: Two Different Expert Aggregation Problems

pert 3 provides information $n_3(E)$ and $n_3(F,B)$. (See Figure 2.) Appendix II shows that the updated subjective probabilities for the three possibilities — using only expert 2 and expert 3 — are:

$$P_E = \frac{1+n_3(E)}{3+n_2(E,F)+n_3(E)+n_2(B)+n_3(F,B)}[1+\frac{n_2(E,F)}{3+n_3(E)+n_2(B)+n_3(F,B)}]$$

$$P_F = \frac{1+n_2(E,F)\frac{2+n_2(B)+n_3(F,B)}{3+n_2(B)+n_3(F,B)+n_3(E)}+n_3(F,B)\frac{2+n_3(E)+n_2(E,F)}{3+n_3(E)+n_2(B)+n_2(E,F)}}{3+n_3(E)+n_2(B)+n_2(E,F)+n_3(F,B)}$$

$$P_B = \frac{1+n_2(B)}{3+n_2(E,F)+n_3(E)+n_2(B)+n_3(F,B)}[1+\frac{n_3(F,B)}{3+n_3(E)+n_2(B)+n_2(F,B)}]$$

Requiring both experts to use a common partition might be especially difficult in this example which could hinder the application of conventional approaches. But since our approach has prespecified a reference partition which distinguishes between events $E$, $F$ and $B$, the decision maker can assign a probability to event $F$ — even though none of the experts includes event $F$ in their partition. As Appendix II notes, this involves first computing an updated probability for $E$ using the information which expert 2 provides for $E$ or $F$ and the information which expert 3 provides for $E$. A similar procedure is used to compute an updated probability for $B$. We then compute the updated probability for $F$ as $1 - p_E - p_B$.

## 4 Summary

A substantial literature describes how to aggregate the probability assessments of different experts. Virtually all of these methods assume that experts define their probabilities over the same partitioning of events. But in reality, different experts often use different categories in describing outcomes for the same event. (This problem may be especially serious with large, heterogeneous expert panels (Hoffman et al., 2007) or in web-based contexts (French et al, 2007).) As a result, it is easier for experts to assess subjective probabilities over the different partitions corresponding to these different categorizations. This leads to the problem of aggregating expert probabilities defined over different partitions.

This paper offers a solution to this problem. This solution will, in general, only be computable algorithmically. But in order to compare our approach to previous approaches — which require that experts all use the same partition — we examined the formula in two simple cases (based on the automotive industry.) In both cases, our approach provides more detailed probability assessments than traditional methods. The second case also provides an example where it could be potentially impossible (or, at least, rather difficult) to apply traditional approaches. As our examples show, it is, in fact, possible for the client to gather information about an event, even if none of the experts directly provides information about that event. In this case, the aggregation procedure could have helped the client discover a useful distinction to make in thinking about the decision problem.

### References


Berger, J. 1985. *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York.

Bernardo, J., A. Smith. 1994. *Bayesian Theory.* Wiley: Chichester.

Billingsley, P. 1995. *Probability & Measure.* John Wiley & Sons, New York.

Clemen, R., R. Winkler. 1999. Combining Probability Distributions From Experts in Risk Analysis. *Risk Analysis* **19,** (2) 187-203.

DeFinetti, B. 1974 *Theory of Probability.* Wiley: Chichester.

Diaconis, P., D. Freedman. 1984. A dozen deFinetti-style results in search of a theory. *Ann.Inst.H. Poincare.* 23, 397-423.

Dickey,J.,J.Jiang, J. Kadane. 1987. Bayesian Methods for Censored Categorical Data. *Journal of the American Statistical Association* **82** (399), 773-781.

Fischhoff, B., Slovic P., S. Lichtenstein (1978). Fault-trees:Sensitivity of Estimated Failure Probabilities to Problem Representation. *Journal* *of Experimental Psychology:Human Perception and Performance,* 4:330.

Fox, C., B. Clemen. 2005. Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science,*51 , 1417-1432.

Fox, C, Y. Rottenstreich. 2003. "Partition Priming in Judgement under Uncertainty." *Psychological Science.*14,3.

Fox, C., K. See. 2003. (Hardman, D & L. Macchi, ed.) *Thinking: Psychological Perspectives on Reason, Judgement & Decision Making.* John Wiley & Sons.

Fox, C., A. Tversky.(1995) "A Belief-Based Account of Decision-making under Uncertainty." In D.Kahneman & A. Tversky. *Choices, Values & Frames.* Cambridge University Press, Cambridge.

French, S. 1985. Group consensus probability distributions: A critical survey. Bayesian Statistics, 2, 183-202.

French, S., D. Insua, F. Ruggeri. 2007. E-Participation and Decision Analysis. *Decision Analysis.* 4 **(4),** 211-226.

Gelman, A, J. Carlin, H. Stern, D. Rubin .1995. *Bayesian Data Analysis.* Chapman & Hall, New York.

Genest, C., Zidek, J.V. 1986. Combining probability distributions: a critique & an annotated bibliography. *Statistical Sciences.*1, 114-48.

Good, I. 1950. *Probability and the Weighting of Evidence.* Hafner, New York.

Hoffman, S., P. Fischbeck, A. Krupnick, M. McWilliams. 2007. Elicitation from Large, Heterogeneous Expert Panels: Using Multiple Uncertainty Measures to Characterize Information Quality for Decision Analysis. *Decision Analysis.* 4 **(2),** 92-109.

Keisler, J., H. Keisler. 1999. The Report Problem. www.math.wisc.edu/ keisler/ReportProb-010108-web.pdf.

D. V. Lindley, A. Tversky, R. V. Brown. 1979. On the Reconciliation of Probability Assessments. *Journal of the Royal Statistical rm Society. Series A (General), 142* **(2)** , 146-180.

Little, R., D. Rubin. 1987. *Statistical Analysis with Missing Data.* Wiley, New York.

Morgan, G., M. Henrion. 1990 *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press, Cambridge.

Mosleh, A., V. Bier. 1996. Uncertainty about Probability: A Reconciliation with the Subjectivist Viewpoint. *IEEE Transactions on Systems,*

*Man and Cybernetics: Part A.* 26,3.
Predd, J., D. Osherson, S. Kulkarni, H. Poor. (2008). Aggregating Probability Forecasts from Incoherent and Abstaining Experts. *Decision Analysis.* in press.
Redelmeier,D.,Kohler,D., Liberman V., A. Tversky. 1995. Probability Judgements in Medicine:Discounting Unspecified Alternatives. *Medical Decision Making.* 15:226.
Rottenstreich, Y., A. Tversky. 1997. "Unpacking,Repacking & Anchoring: Advances in Support Theory." *Psychological Review.*2, 406-415.
Rowe, Wright. 2001. Expert Opinions in Forecasting. Role of the Delphi Technique. In: Armstrong (Ed.): *Principles of Forecasting: A Handbook of Researchers and Practitioners,* Kluwer Academic Publishers, Boston.
Schlaiffer, J. 1997. "Analysis of Incomplete Multivariate Data." *The CRC Press.*
Wallsten, T., D. Budescu. 1980. Encoding Subjective Probabilities: A Psychological and Psychometric Review. *EPA Research Report.*
Wong, T. 1998. Generalized Dirichlet distribution in Bayesian analysis. *Applied Mathematics and Computation*, 97, 165-181

## APPENDIX I

Suppose one expert's partition is the atomic partition $\{E, F, B\}$ while another expert's partition is $\{(E, F), B\}$. Then if $K$ is some positive integer, the subjective probability (or the mean value) is given by

$$\frac{\int_p p_E p_E^a p_F^b p_B^c (p_E + p_F)^K}{\int_p p_E^a p_F^b p_B^c (p_E + p_F)^K}$$

for some values $a, b, c$, which has the form

$$\int_p \frac{\sum_{j=0} \binom{K}{j} p_E^{a+j+1} p_F^{b+K-j} p_B^c}{\int_p \sum_{j=0} \binom{K}{j} p_E^{a+j} p_F^{b+K-j} p_B^c}$$

Define

$$E[p_E|j] = \frac{\int_p \binom{K}{j} p_E^{1+a+j} p_F^{b+K-j} p_B^c}{\int_p \binom{K}{j} p_E^{a+j} p_F^{b+K-j} p_B^c}$$

which is the mean of $q_E$ if we knew that $j$ of the $K$ $(E, F)$ responses were, in fact, response $E$. But

$$E[p_E|j] = \frac{a+j+1}{a+b+K+c+3}$$

Also define

$$w_j = \frac{\int_p \binom{K}{j} p_E^{1+a+j} p_F^{b+K-j} p_B^c}{\int_{p_E} p_E^a p_F^b p_B^c (p_E + p_F)^K} = \frac{\int_p \binom{K}{j} p_E^{a+j} p_F^{b+K-j} p_B^c}{\sum_{j=0}^K \int_p \binom{K}{j} p_E^{a+j} p_F^{b+K-j} p_B^c}$$

But

$$w_j \propto \binom{K}{j} \frac{\Gamma(a+j+1)\Gamma(b+K-j+1)\Gamma(c+1)}{\Gamma(a+b+K+c+3)} \propto \frac{\binom{K}{j}}{\binom{a+b+K+2}{a+j+1}} \propto \frac{\binom{K}{j}\binom{a+b+2}{a+1}}{\binom{a+b+K+2}{a+j+1}}$$

which is proportional to a beta binomial distribution with parameters $a+1, b+1, K$. Since the mean of $j$ where $j$ has a beta-binomial distribution is $K\frac{a+1}{a+b+2}$, we get the mean of $p_E$ as

$$\begin{aligned} \sum_j w_j \frac{a+j+1}{a+b+K+c+d+4} &= \frac{a + \frac{K(a+1)}{a+b+2} + 1}{a+b+K+c+3} \\ &= \frac{a+1}{a+b+c+K+3} + \frac{K}{a+b+c+K+3}\frac{a+1}{a+b+2} \end{aligned}$$

We can use symmetry to infer $E[p_F]$ and deduce $E[p_B] = 1 - E[p_E] - E[p_F]$.

## APPENDIX II

We first focus on computing the subjective probability for the second expert. Using the argument of Appendix I leads to a density proportional to

$$(p_E + p_F)^K p_3^c p_E^a (p_F + p_B)^J$$

If $K$ and $J$ are integral, we can write this as

$$\sum_{i=0,j=0}^{K,J} \binom{K}{i}\binom{J}{j} p_E^{K-i+a} p_F^{i+j} p_B^{c+J-j}$$

with the mean of $p_E$ being

$$\frac{\sum_{i=0,j=0}^{K,J} \binom{K}{i}\binom{J}{j} p_E^{1+K-i+a} p_F^{i+j} p_B^{c+J-j}}{\sum_{i=0,j=0}^{K,J} \binom{K}{i}\binom{J}{j} p_E^{K-i+a} p_F^{i+j} p_B^{c+J-j}}$$

As before, we define

$$E[q_E|ij] = \frac{\int_p \binom{K}{i}\binom{J}{j} E_1^{K+1-i+a} p_F^{i+j} p_B^{c+J-j}}{\int_p \binom{K}{i}\binom{J}{j} p_E^{K-i+a} p_F^{i+a} p_B^{c+J-j}}$$

and also define

$$w_{ij} = \frac{\int_p \binom{K}{i}\binom{J}{j} p_E^{K+a-i} p_F^{i+j} p_B^{c+J-j}}{\sum_{ij}\int_p \binom{K}{i}\binom{J}{j} p_E^{K+a-i} p_F^{i+j} p_B^{c+J-j}}$$

Thus we can write the mean of $p_E$ as $\sum_{ij} E[p_E|ij]w_{ij}$. But

$$E[p_E|ij] = \frac{K+a+1-i}{K+a+c+J+3} = \frac{K+a+1}{K+a+c+J+3} - \frac{i}{K+a+c+J+3}$$

Also

$$w_i = \sum_j w_{ij} = \frac{\binom{K}{i}\sum_j \binom{J}{j}\Gamma(K+a-i+1)\Gamma(i+j+1)\Gamma(c+J-j+1)}{\sum_i \binom{K}{i}\sum_j \binom{J}{j}\Gamma(K+a-i+1)\Gamma(i+j+1)\Gamma(c+J-j+1)}$$

But

$$\sum_j \binom{J}{j}\Gamma(i+j+1)\Gamma(c+J-j+1) = \Gamma(i+c+J+2)$$

so that

$$w_i = \frac{\binom{K}{i}\Gamma(K+a-i+1)\Gamma(c+J+i+2)}{\sum_i \binom{K}{i}\Gamma(K+a-i+1)\Gamma(c+J+i+2)} \propto \frac{\binom{K}{i}\binom{c+a+J+2}{c+J+2}}{\binom{K+c+a+J+3}{c+J+i+2}}$$

which is another beta-binomial distribution with mean $K\frac{c+J+2}{a+c+J+3}$. Thus the mean of $p_E$ is

$$\frac{K+a+1}{K+a+c+J+3} - \frac{\sum_i i\ w_i}{K+a+c+J+3} = \frac{a+1}{K+a+c+J+3} + \frac{K}{K+a+c+J+3}\frac{a+1}{a+c+J+3}$$

This can be written more simply as

$$\frac{a+1}{K+a+c+J+3}[1+\frac{K}{a+c+J+3}]$$

From symmetry, the mean of $p_B$ is

$$\frac{c+1}{K+a+c+J+3}[1+\frac{J}{a+c+K+3}]$$

The mean $p_F$ can be computed from $1-E[p_E]-E[p_B]$ which gives

$$\frac{1+K\frac{c+J+2}{c+J+a+3}+J\frac{a+K+2}{a+c+K+3}}{a+c+K+J+3}$$