Learning from the Probability Assertions of Experts

Mario Di Bacco
University of Bologna
Patrizio Frederic
University of Bologna
Frank Lad
University of Canterbury*

Abstract

We develop a new outlook on the use of experts' probabilities for inference, distinguishing the information content available to the experts from their probability assertions based on that information. Considered as functions of the data, the experts' assessment functions provide statistics relevant to the event of interest. This allows us to specify a flexible combining function that represents a posterior probability of interest conditioned on all the information available to any of the experts; but it is computed as a function of their probability assertions. We work here in the restricted case of two experts, but the results are extendible in a variety of ways. Their probability assertions are shown to be almost sufficient for the direct specification of the desired posterior probability. A mixture distribution structure that allows integration in one dimension is required to yield the complete computation, accounting for the insufficiency. One sidelight of this development is a display of the moment structure of the logitnormal family of distributions. Another is a generalisation of the factorisation property of probabilities for the product of independent events, allowing a parametric characterisation of distributions which orders degrees of dependency. Three numerical examples portray an interesting array of combining functions. The coherent posterior probability for the event conditioned on the experts' two probabilities does *not* specify an externally Bayesian operator on their probabilities. However, we identify a natural condition under which the contours of asserted probability pairs supporting identical inferences are the same as the contours specified by EB operators. Our discussion provides motivation for the differing function values on the contours. The unanimity and compromise properties of these functions are characterised numerically and geometrically. The results are quite promising for representing a vast array of attitudes toward experts, and for empirical studies.

Key Words: Combining probabilities, expert judgement, unanimity, compromise, externally Bayesian pooling operators, exchangeability, conditional independence, logitnormal distribution.

1 Introduction

Resolving statistical problems of combining information presented by two experts hinges on the difficult feature that we typically regard the opinions of the experts dependently. That is, knowledge of what one expert says tells us something about what the other expert will say ... even when the second expert does not know specifically what the first expert has said. The dependence in our attitude towards the experts arises from recognition of the similar training and apprenticeship they would have engaged in order to become recognised as experts. Deciphering the information that they share from their unique contributions is a statistical problem amenable to probabilistic analysis.

*Address for correspondence: Frank Lad, Department of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand. Email: F.Lad@math.canterbury.ac.nz

We propose a new formal resolution to the problem of combining the information presented via probability assertions of two experts. We presume that "the statistician" is a coherent expert on inferential logic and a generally educated person, but is *not* an expert on the specific subject matter at hand. The statistician may be consulting for an interested client who is not an expert on the subject matter either. Even when the statistician (or client who is represented) cannot imagine the *content* of the expertise to be elicited from the two experts, he/she may be able to formulate an uncertain opinion regarding the *strength* of the evidence they might be able to provide for the question at hand. Such judgements motivate the reliance on an expert in the first place, as well as the decision to consult a second expert for an opinion. What would be unknown is whether the strength of the information is supportive or in opposition to the occurence of the event in question. We shall describe how the consulting statistician's uncertainty may be expressed and the form in which the experts' probability assertions must be combined in order to make coherent inferences from their assertion. Although each of their probilities is presumed to be a sufficient summary of their own information sources, the two probabilities are not sufficient for all of the information jointly. We identify conditions under which the product of the experts' odds ratios is a statistic that is almost sufficient to portray the information both of their expertise provides, and we show what assessments need to be made by the statistician to complete the combination function via a mixture intergration. Under more general conditions, sufficiency requires the multiplicand odds ratios to be raised to differing powers. However, in no case are the experts' probabilities completely sufficient for all the information available to the experts.

The question of combining experts' probabilities to inform one's uncertainty about an event has received regular attention over the past forty years. Nonetheless, the developments we present here are novel. An insightful technical review of research relevant to our contribution appears in an article by Clemen and Winkler (1999), and a broader review of related literature appears in a book by French and Rios Insua (2000). We shall recall here only two themes of the multi-faceted research on this problem that relate specifically to our work.

One line of research stemming from the original work of McConway (1981) has developed important results on the mechanical pooling of individual opinion distributions into a consensus opinion. In this tradition, functional analysis is used to identify all pooling functions that satisfy various reasonable sounding properties. Once proposed among desireable features of a combination function is that the pooled distribution, when updated by a likelihood function agreed upon by all contributing individuals, should equate with the pooled distribution generated from the individuals' updated distributions. This feature was termed "external Bayesianity" by Madansky (1964). A complete array of studied properties appears in the review of French and Rios Insua (4.5-4.10, pp. 111-115). The promise of this line of research was exemplified by a synthetic technical result characterising the algebraic form of externally Bayesian pooling operators, achieved by Genest, McConway and Schervish (1986). In the context of a single event, H, assessed by two experts with the probabilities p_1 and p_2 (the context we shall study here), it says that a pooling function $g(p_1, p_2)$ yielding a "combined probability" is an externally Bayesian operator if and only if it has the form

$$g(p_1, p_2) \;=\; p_1^w p_2^{(1-w)} / \; [p_1^w p_2^{(1-w)} + (1-p_1)^w (1-p_2)^{(1-w)}] \quad \text{for some} \;\; w \;\in\; [0,1] \; . \quad (1)$$

Two other possible properties of combining functions have been aired as the unanimity principle, proposed in Morris (1974), and its extension to the compromise principle, studied in Clemen and Winkler (1990). The former requires that when the two experts assert the same probability value then the combination function agrees as well: $g(p,p) = p$. The latter requires more generally that the combination functions yields a value between the two experts' probabilities when they do not agree: $g(p_1 < p_2) \in [p_1, p_2]$. When studying an array of combination functions that do not universally honour these principles, Clemen and Winkler assess the extent to which they do and do not. For example, it is evident from (1) that external Bayesianity would imply unanimity. Further developments along the line of establishing families of distributions that exhibit specific properties appear in the article of Dawid, DeGroot and Morterra (1995) followed by extensive discussion.

A second tack was originated by Winkler (1968) and Morris (1974), followed by works of Lindley, Tversky and Brown (1979), Good (1979) and Dickey (1980). It explicitly treats experts' probabilities as data relevant to an event, and derives coherent forecasting distributions implied by various assumptions about the generation of this "data." Lindley's further considerations (1985) were the first to claim that the property of external Bayesianity is not appropriate to a coherent inference from experts' assertions. Algebraic details of his analysis rely on the use of a multivariate normal distribution over the experts' log-probabilities, applied only to contrasts. With this proviso, O'Hagan concurred in the discussion, though Bernardo questioned it.

The outlook we present here follows Winkler's tack of treating experts' probabilities as statistical data relevant to the uncertain event at hand. But we highlight the fact that these probabilities are statistics, that is, functions of the complete data available to the experts. We investigate the sufficiency of these statistics for the complete information available to the experts that motivates their probabilities. In deriving an exact solution, we find that the coherent posterior probability does not specify an externally Bayesian operator; nor does it even support unanimity of the inference with the two experts when they agree in their probability assertion value. This non-concurrence arises even in the context that the experts are highly trusted, a context that we define precisely. Moreover, we can identify precisely the extent to which the Clemen-Winkler extension to the compromise principle will be honoured. When algebraic details are addressed, a logitnormal rather than a lognormal distribution over experts' probabilities is motivated.

An important feature of our analysis is the formal distinction between information that the two experts hold in common and the information they hold only separately. This distinction has been made as early as the works of Zeckhauser (1971) and Winkler (1981), and is studied in clever examples by Clemen (1987) yielding interesting insights when formal conditions of exchangeable quantities are involved. We analyse this distinction in a more general context and show that its recognition is useful for the problem of decyphering the evidential impacts of the distinct information types.

The survey of French and Rios Insua references several applied studies using experts' judgements as data. A further application by Lad and Di Bacco (2002) exemplifies a distribution which assesses a pair of experts' probabilities exchangeably, but subsequently learns not to regard succeeding pairs of their assertions exchangeably. The authors' reconsideration of that problem led to the analysis reported in the present article. While the computational results presented in the present article are merely numerical, they do exemplify very general features of the problem that are worth examining. The easiest first application of our results to the data presented in Clemen and Winkler (1990) is currently under investigation.

2 Trusting and Cooperative Scientific Inference

At the heart of the problem we address here is an event whose value is unknown to anyone. We shall denote it by H. For example, H might equal 1 if a patient has a cancer of the lung, and equal 0 if not. Again, H might equal 1 if a skull discovered in an anthropological excavation is the skull of a male, and equal 0 if from a female; or H might equal 1 if a scoop of about 50 honey bees taken from a beehive shows evidence of Varroa mites in a powdered sugar test, and equal 0 if not.

The problem revolves upon the uncertain knowledge of three different people: a statistician, whom we shall designate by S, for example in denoting probability assertions by $P_S(\cdot)$; a person who is considered by S to be a specialist expert about matters related to H, whose probability assertions we shall denote by $P_1(\cdot)$; and a second expert on such matters, whose assertions we denote similarly by $P_2(\cdot)$. The statistician is typically a consultant for still another person who has a practical interest in H but is no more specifically informed about H than is the statistician. However, in our notation we attribute the opinions to S.

We also distinguish three different sources of information relevant to this problem. There is basic background information that would be known by most educated people, including the statistician. We denote this basic information by the event B. (Following the style of de Finetti, we do not distinguish between an event and its numerical indicator. An event is defined to equal 1 if the statement that defines it is true, and to equal 0 if it is false.) Of course B itself might equal 1 or 0, because these facts might in principle be true or false. However, most people with very basic background knowledge of the general subject would know it to equal 1. For example, most people know that in human populations the sex ratio of males to females at birth is almost equal to but is slightly larger than 1. In principle, it might not even be close to 1, and in populations of different species the sex ratio at birth is not close to 1. However, most people know that in human populations it is. Since all three people involved in our problem are presumed to be aware of the truth of B, there will be no need for us to declare it specifically in any probability assertions we shall discuss. It can always be presumed to be a motivating component of any probability assertion $P(\cdot)$, however P may be subscripted. Nonetheless, B needs to be recognised as relevant, and is mentioned explicitly in some of our discussion.

A second type of information concerning our problem is information that would be known by most any specialist expert, but would generally not be known or even thought about by people who are not specifically knowledgeable experts. Again, in principle this information might or might not be true. For example, most beekeepers (with the specialist knowledge they gain) would know that in a healthy beehive, the sex ratio of male to female bees at birth in the early spring is close to 1/25. This information is not widely known among the generally educated public, including the statistician who would be uncertain about the ratio if asked, and who may never have even thought about the role this ratio plays in the health of a hive. We shall denote such information by the event F.

A third type of information is detailed information that is known uniquely by specifically informed experts. The information known specifically to expert 1 in our problem is denoted by G_1, and that specifically known to expert 2 by G_2. For example, expert 1 might know that a specific beehive has been treated with the insecticide "Capstan" during the previous year, although this is unknown to expert 2. Alternately, expert 2 might know that the "mite count" in this very same beehive during a recent check of the bottom board screen was 3, while this is unknown to expert 1.

When asked to assert their probabilities for event H, we shall suppose that these three individuals respond with the numbers

$$P_S(H) \ = \ \pi_S, \qquad P_1(H) \ = \ \pi_1, \qquad \text{and} \qquad P_2(H) \ = \ \pi_2 \quad . \tag{2}$$

3 How should the statistician use the probability assertions of the experts?

The problem we address here does *not* concern the very real possibility that two different people may assert different probabilities for an event even when informed by the very same data. We shall assume the statistician trusts the two experts to the extent he/she presumes that if informed by the same training, experience and information as the experts, the statistician would assert the same probabilities they do. The reason for relying on the experts' assertions is that not only does the statistician lack the training and general experiences of these experts along with their specific information, but the statistician cannot even imagine what the relevant information possibilities $F, \tilde{F}, G_1, \tilde{G_1}, G_2, \tilde{G_2}$ are (meaning, how they are defined) so to assert personal probabilities for H conditioned upon them!

Nonetheless, the statistician is aware that in asserting their personal probabilities $P_1(H)$ and $P_2(H)$, these experts are making use of their knowledge of FG_1 and FG_2 in assessing their probabilities. Although unaware of what this information and its relevance to the experts' probability assertions might be, the statistician would like very much to use both of their their announced probabilities to specify his/her own appropriate $P_S(H|FG_1G_2)$ based on the information available to both of them. We shall derive explicitly how this can be done using only the probabilities elicited from the experts, π_1 and π_2, rather than requiring a complete description of the information on which they base their assertions.

4 Modelling the experts' probabilities as statistics

Firstly, the statistician realises that each expert i's probability assertion $P_i(H)$ is identical to $P_i(H|FG_i)$. This is due to the theorem of total probability because expert i also asserts $P_i(FG_i) = 1$; for expert i knows the information defining FG_i to be the case. Moreover, as we explained above, S trusts each expert i to the extent that $P_S(H|FG_i)$ would also equal each $P_i(H|FG_i)$ if only S had the training, experience and information to know what these conditioning quantities were. Now, being told the values of these probabilities by the experts, S agrees with them and thus is able to assert them too! To summarise this realisation formally, we can write

$$P_S(H|FG_i) \;=\; P_S(H|FG_i, \pi_i) \;=\; P_S(H|\pi_i) \;=\; \pi_i \;=\; P_i(H|FG_i) \;=\; P_i(H) \;\; \text{for each } i \;. \tag{3}$$

It is worth a technical remark here that equation (3) amounts to the specification that expert i's probability is consider by S to be *sufficient* for the information on which it is based: $P_S(H|FG_i, \pi_i) \;=\; \pi_i$. That is, the data provided by the expert's π_i is sufficient to express the conditional probability given a complete description of the information upon which it is based, FG_i . Although the statistician would agree with either expert's assertion, asserted on its own, the question remains how to use the information provided by *both* of their assertions. What we need to develop is the appropriate value desired by the statistician, $P_S(H|FG_1G_2)$. This is a probability that conditions on all the relevant information, whether shared or unique to the two individuals.

We presume a second useful agreement to feature in this scenario of cooperative and trusting science. Both experts are presumed to understand the same evidential content of their common information, F, for the unknown event, H. That is, they would agree on an asserted probabilistic basis for the likelihoods

$$P_i(F|H) \equiv \mathcal{L}(H;F) \quad \text{and} \quad P_i(F|\tilde{H}) \equiv \mathcal{L}(\tilde{H};F) \quad . \tag{4}$$

(We use tilde notation, as in $\tilde{H}$, to denote the complement of an event throughout this paper.) It is not required that these likelihood probabilities are evaluated or asserted by the experts. In fact, we presume explicitly that they do not announce such valuations. Nonetheless, were they able to compare notes with one another, we presume they would agree in their assessments.

Using the same notation of likelihood symbol, $\mathcal{L}$, we shall denote the specific likelihoods defined by the considered opinions of experts 1 and 2 about their personal information sources by

$$P_i(G_i|H) \equiv \mathcal{L}(H;G_i) \quad \text{and} \quad P_i(G_i|\tilde{H}) \equiv \mathcal{L}(\tilde{H};G_i) \quad . \tag{5}$$

Moreover, we presume that S would agree with these likelihoods as well if the probability assertions defining them were publicly specified.

These agreements that we presume among the three people are natural to presume about the relations among an educated person and the probabilistic assertions of two trusted experts. However, they explicitly preclude situations where the statistician may distrust an "expert's" assessments on account of suspected bias as may be the case in an adversarial situation such as legal evidence presented by an opponent. The overall framework we are developing may be expanded to address such situations, but that is a matter beyond our considerations here. Our framework does not preclude the experts having different approaches or personalities underlying their investigations. In fact it expressly allows for this. That is why a second expert is consulted, for another perspective on the matter of H. What we do presume is that the statistician does not have specific knowledge of the experts' differences ... only knowing that they are both recognised experts worth consulting. Nonetheless, the general result we shall derive allows for this as well, that the specific information of one expert may be considered to preempt that of the other to some extent.

Before continuing with the analysis of how the experts' announced probabilities are now to be used, let us introduce a bit more useful notation. When the experts announce their probabilities, it will be convenient for us to transform them into odds ratios (in favour), and to use the notation

$$\rho \equiv \frac{\pi_S}{1-\pi_S} \quad \text{and} \quad T_i \equiv \frac{P_i(H|FG_i)}{P_i(\tilde{H}|FG_i)} = \frac{\pi_i}{1-\pi_i} \equiv t_i \quad for\ i=1,2 \ . \tag{6}$$

Similarly, we transform the likelihoods into likelihood ratios, denoted by

$$\ell_F \equiv \frac{\mathcal{L}(H;F)}{\mathcal{L}(\tilde{H};F)} \quad \text{and} \quad \ell_{G_i} \equiv \frac{\mathcal{L}(H;G_i)}{\mathcal{L}(\tilde{H};G_i)} \quad for\ i=1,2 \ . \tag{7}$$

Notice one difference between the quantities T_i and ℓ_F, ℓ_{G_i}: the quantities T_i are public quantities, reported to the statistician by the experts; whereas the quantities ℓ_F and ℓ_{G_i} have not been reported publicly as such. They have merely been assessed either formally or informally by the experts when assessing the probabilities $P_i(H)$ that they do report. Of course ρ is a number asserted by S via $P_S(H)$. We will use all of these quantities in the derivation that follows.

A second notice regarding the quantities T_i is in order. Before the statistician asks the experts for their probability assertions regarding H, the numbers they will respond are unknown. The statistician, of course, has no way of knowing the content of these assertions, but is merely uncertain about what the experts might say. It is for this reason that the upper case quantities T_i are defined in (6) as unknown variables, while the lower case t_i denote the specific values these quantities take on as announced by the experts. We shall see the relevance of this feature later on in our development. The same comment could be made about the likelihood ratios that we have denoted by ℓ_F, ℓ_{G_1} and ℓ_{G_2}. Although the experts will never announce the values of these quantities in our scenario, a consideration of what their values might be will be found relevant to the statistician's inferences from the probabilities the experts do announce.

In this context, it is worth noting now that the presumptions of equation (3) imply that the statistician's prevision (expectation) for the expert's probability assertion value must be identical to the statistician's own (relatively uninformed ... only by B) probability. For

$$P_S[P_i(H|FG_i)] = P_S[P_S(H|FG_i)] = P_S(H) \quad . \tag{8}$$

Readers who are unfamiliar with de Finetti's language of "prevision" would know this result by the familiar statement that "the expectation of a conditional expectation must equal the unconditional expectation."

5 Computational representation of $P_S(H|FG_1G_2)$

Our goal in this section is to analyse the statistician's posterior probability for H given all the information available to the two experts, $P_S(H|FG_1G_2)$, expressed in terms of the probability for H conditioned only on the experts' probability assertions, viz.,

$$P_S[H|FG_1G_2) = P_S(H|(P_1(H|FG_1)=\pi_1)\ (P_2(H|FG_2)=\pi_2)] \ .$$

It will be convenient to derive the computational form of this probability in terms of the posterior odds ratio in favour of H:

$$Odds_S(H|FG_1G_2) = P_S(H|FG_1G_2)\ /\ P_S(\tilde{H}|FG_1G_2) \ .$$

In deriving this odds ratio, it will be helpful to recall and use a simple relation between any conditional odds ratio and the associated inverse probability ratio, which you can derive for yourself, viz.,

$$Odds(A|B) = \frac{P(B|A)}{P(B|\tilde{A})}\ Odds(A) \tag{9}$$

Applying the general structure of (9) to arguments in the first and third lines that follow, along with standard factoring in the second line, our desired odds ratio reduces to

$$Odds_S(H|FG_1G_2) = \frac{P(G_1G_2F|H)}{P(G_1G_2F|\tilde{H})}\ Odds_S(H)$$

$$
\begin{aligned}
&= \frac{P(G_2|G_1FH)}{P(G_2|G_1F\tilde{H})}\ \frac{P(G_1|FH)}{P(G_1|F\tilde{H})}\ \frac{P(F|H)}{P(F|\tilde{H})}\ Odds_S(H) \\
&= \frac{P(G_2|G_1FH)}{P(G_2|G_1F\tilde{H})}\ \frac{Odds(H|G_1F)}{Odds(H|F)}\ \frac{P(F|H)}{P(F|\tilde{H})}\ Odds_S(H) \\
&= \frac{P(G_2|G_1FH)}{P(G_2|G_1F\tilde{H})}\ \frac{T_1}{\ell_F\ \rho}\ \ell_F\ \ \rho\ \ . \qquad (10)
\end{aligned}
$$

The final line uses the notation we defined in (6) and (7), and allows for some cancellations. The conditional probability ratio $\frac{P(G_2|G_1FH)}{P(G_2|G_1F\tilde{H})}$ will require some discussion.

5.1 An array of expected relations among the information contents of F, G_1 and G_2

To complete the analysis of (10) we need now to address the statistician's attitude toward the relation between the information available uniquely to each of the experts, G_1 and G_2, and F and H. This relation arises in the conditional probabilities in the numerator and denominator of (10). In terms of the numerator for example, the question rests as to whether knowing H is the case and knowing all the information F that is available to any "recognised expert", the information provided by knowing additionally what expert 1 knows would provide further information about the truth of what expert 2 knows. Should $P(G_2|G_1FH)$ simply equal $P(G_2|FH)$? Or should $P(G_2|G_1FH)$ be augmented on account of the additional conditioning on G_1? The simplest, and we believe often relevant assertion would be to regard the additional information sources G_1 and G_2 exchangeably via a mixture of conditional independent assertions given FH and given $F\tilde{H}$. Such an assertion is appropriate in situations where much more information about H is required, above and beyond the general information of experts provided by F, and it is expected that each of the separate experts would have extensive individual information to provide, unrelated to that available to the other. Intuitively, one's feelings in these situations are much like those experienced at the early stages of a huge picture puzzle. Technically it would mean that

$$
\begin{aligned}
P_S(G_2G_1|FH) &= P_S(G_1|FH)\ P_S(G_2|FH)\ , \quad \text{and} \\
P_S(G_2G_1|F\tilde{H}) &= P_S(G_1|F\tilde{H})\ P_S(G_2|F\tilde{H}) \qquad (11)
\end{aligned}
$$

or equivalently,

$$
P_S(G_2|G_1FH) = P_S(G_2|FH)\ , \quad \text{and} \quad P_S(G_2|G_1F\tilde{H}) = P_S(G_2|F\tilde{H})\ . \qquad (12)
$$

However we can extend this analysis to a wider class of exchangeable and non-exchangeable assessments which would be appropriate in situations where it is felt that we are accumulating the final bits of information that could possibly be available about H. This is signalled by the feeling that "things are finally falling into place", much as at the final stages of completing a picture puzzle. In this case, the evidence provided by G_1 may be felt to make much more likely the evidence provided via G_2, and perhaps vice-versa in a symmetric fashion (or perhaps not symmetrically if G_1 is expected to be more informative than G_2). Such assertions can be expressed technically by generalising the assertions of equations (11) and (12) to

$$
\begin{aligned}
P_S(G_2G_1|FH) &= P_S(G_1|FH)^{\alpha}\ P_S(G_2|FH)^{\beta}\ , \quad \text{and} \\
P_S(G_2G_1|F\tilde{H}) &= P_S(G_1|FH)^{\alpha}\ P_S(G_2|F\tilde{H})^{\beta} \qquad (13)
\end{aligned}
$$

for appropriate values of α and β, or equivalently

$$
\begin{aligned}
P_S(G_2|G_1FH) &= P_S(G_1|FH)^{\alpha-1}\ P_S(G_2|FH)^{\beta}\ , \quad \text{and} \\
P_S(G_2|G_1F\tilde{H}) &= P_S(G_1|F\tilde{H})^{\alpha-1}\ P_S(G_2|F\tilde{H})^{\beta}\ . \qquad (14)
\end{aligned}
$$

A complete explanation of the allusion to "appropriate values of α and β" is deferred a separate tech report by Frank Lad, entitled "Factoring Bivariate Distributions." It is currently available to any interested reader as a .pdf file. Please email him if you are reading this draft and are interested. For now, notice firstly that when $\alpha = \beta = 1$, assertions (13) and (14) reduce to assertions (11) and (12) which are a special case. Exchangeable attitudes toward G_1 and G_2 are allowed (though not required) only by (α, β) pairs within the unit-square whose sum is not less than 1. Surprisingly, the generalisation of the well-known independence equation $P(AB) = P(A)P(B)$ to $P(AB) = P(A)^\alpha P(B)^\beta$ has never been studied heretofore. The complete parameter space of coherent (α, β) pairs is somewhat unusual, as described in the appendix. For now, we mention only that every infinitely exchangeably extendible distribution over two events subscribes to the restrictions of (13) and (14) for some parameter configuration in the region $\alpha = \beta \in [\frac{1}{2}, 1]$.

Using (13) and (14) we can represent virtually every interesting assessment of the experts' individual information sources. Thus, we shall now continue our development of $P_S(H|FG_1G_2)$ through equation (10) using this representation. However once we have derived the desired result, we shall continue our numerical examples using only the special case of conditional independence of the experts' information sources, identified by $\alpha = \beta = 1$.

5.2 Experts' probabilities as 'almost sufficient' statistics for FG_1G_2

Inserting the stipulations of (13) into (10) and performing the allowable cancellations yields the result

$$\begin{aligned} Odds(H|FG_1G_2) &= \frac{P(G_1|FH)^{\alpha-1}\ P(G_2|FH)^\beta}{P(G_1|F\tilde{H})^{\alpha-1}\ P(G_2|F\tilde{H})^\beta}\ T_1 \\ &= \left(\frac{Odds(H|FG_1)}{Odds(H|F)}\right)^{\alpha-1} \left(\frac{Odds(H|FG_2)}{Odds(H|F)}\right)^{\beta} T_1 \\ &= \frac{T_1^\alpha\ T_2^\beta}{(\rho\ \ell_F)^{\alpha+\beta-1}} \quad . \end{aligned} \tag{15}$$

The second equality derives from an application of (9), with the final line following from notational substitutions and appropriate algebra.

In the special case of assessed conditional independence regarding G_1 and G_2 given FH and $F\tilde{H}$, this reduces to

$$Odds(H|FG_1G_2) = \frac{T_1\ T_2}{\rho\ \ell_F} \quad . \tag{16}$$

We can now recognise from (15) that T_1 and T_2 together are *not sufficient* for the complete information in FG_1G_2, even though T_1 is sufficient for FG_1 and T_2 is sufficient for FG_2, as we had noticed after equation (3). Along with ρ, a third statistic would be required to suffice for all three components of this information, viz. the non-elicited value of ℓ_F. Of course ρ has been specified via the statistician's assertion of $P_S(H)$.

In fact, having stipulated the values of α and β, it is the product $T_1^\alpha\ T_2^\beta$ that is almost sufficient for FG_1G_2, requiring only the additional assertion of $\rho\ell_F$ for sufficiency. This statement flows more simply *in the special case* of asserted conditional independence of G_1 and G_2 given FH and given $F\tilde{H}$: *the product of the odds ratios asserted by the two experts is almost sufficient for the totality of the information they provide regarding the event of interest.* In this form, the statement of this result is striking, for the common response of people who hear of the problem of combining probabilities from two experts is to suggest combining them via their arithmetic average, which is *not* a sufficient statistic. The near-sufficiency of the product of the odds ratio for the unobserved (by S) and perhaps not even described (by the experts) informational events F, G_1 and G_2 is a pleasing result. It is possible and even likely that the experts cannot precisely explain the informational substance of their expertise, but only the measure of their resulting inferences from it, their assertion values π_i. It is thus quite pleasing to find that their numerical probability assertions $P_i(H) = \pi_i$ which determine the $T_i = t_i$ are *almost* sufficient to inform the statistician of the inductive content of most everything they know that is relevant to H!

From this point on in this article, we shall restrict the details of our analysis to the special case of the presumed conditional independence of G_1 *and* G_2 *given* FH. All details can be generalised tractably to the more general case. However, until we have a chance to describe the opinion structures represented by values of α and β that are different from 1 beyond the concise presentation appearing in the appendix, generalisations would be superfluous. As we shall see, the special case alone yields very interesting computational results, and its practical relevance is not so very limited. Its strength comes from the realisation that conditioning on H (or $\tilde{H}$) is a strong property, for the knowledge of H gives fairly precise meaning to the information contained in the signal of either G_i.

5.3 Transforming posterior odds to posterior probabilities

We can transform a posterior odds ratio into a posterior probability using the inverse of the odds transformation, $prob = odds/(odds{+}1)$. Applying this inversion to (16) and simplifying the result yields

$$P_S(H|FG_1G_2) \;=\; \frac{T_1T_2}{T_1T_2 + \rho\ell_F} \quad . \tag{17}$$

This useful equation requires some consideration. For until we learn the values of T_1 and T_2 from the experts, we can only imagine what they might be. Furthermore, the value of ℓ_F is never elicited from the experts, so the statistician will always be limited to living with uncertainty about it. Once we learn the values of $T_1 = t_1$ and $T_2 = t_2$, we would be able to derive the desired posterior probability $P_S(H|FG_1G_2)$ as the prevision (expectation) of (17) considered as a conditional prevision, $P_S(H|(T_1 = t_1)(T_2 = t_2)\rho\ell_F)$, viz.

$$P_S(H|(T_1 = t_1)(T_2 = t_2)) \;=\; \int_0^\infty \frac{t_1t_2}{t_1t_2 + \rho\ell_F}\, dF(\rho\ell_F|(T_1 = t_1)(T_2 = t_2)) \quad . \tag{18}$$

We need only to formulate an appropriate form of opinion about ℓ_F, T_1 and T_2.

Remember that the value of ℓ_F has not been elicited from the experts, nor have they even necessarily assessed it explicitly. Nonetheless, ℓ_F is a quantity that a statistician can think about, and perhaps assess a mixing distribution that would allow integration of the conditional probability and the mixing distribution $F(\rho\ell_F|(T_1 = t_1)(T_2 = t_2)$ as required by (18). Realising that the content of ℓ_F is a component of the experts' considerations when asserting their T_1 and T_2, whether considered explicitly or not, we shall investigate next the practicality of assessing this conditional mixing distribution.

6 Assessing a distribution for $\rho\ell_F$ given T_1 and T_2

We shall begin by addressing an assessment of $\rho\ell_F$ jointly with T_1T_2. To do this, we need to analyse S's opinions regarding all the information sources we have delineated in the forms of $\ell_F, \ell_{G_1}, \ell_{G_2}$ and T_1T_2. This will allow us to formalise the conditional mixing distribution that we require. Although the statistician cannot imagine the *content* of the experts' information sources F, G_1 and G_2, he/she will surely have some idea about the *amount* of information these sources contain regarding H. That is what motivates the statistician to ask the experts their opinions about H in the first place. Expecting a large amount of information would mean expecting a probability assertion close to 1 or close to 0. Expecting only a little information would mean expecting an assertion either a little above or a little below the assertion of a generally educated person, π_S.

We are fortunate to be able to transform the representation of $\rho\ell_F$ via the equations

$$\rho\ell_F = \frac{P(H)P_i(F|H)}{P(\tilde{H})P_i(F|\tilde{H})} = \frac{P_i(HF)}{P_i(\tilde{H}F)} = \frac{P_i(H|F)}{P_i(\tilde{H}|F)} \equiv \phi_F \quad , \tag{19}$$

which derive from the Bayes' factorisation. The variable denoted by ϕ_F is subscripted with an F to distinguish it from two other quantities we shall define shortly in a similar way as ϕ_1 and ϕ_2.

The practical merit of transformation (19) is that it would be much simpler for a statistician to assess a mixing distribution for the odds defined as ϕ_F, which are specified in terms of $P_i(H|F)$, than for the likelihood ratio defined by ℓ_F, specified in terms of $P_i(F|H)$ and $P_i(F|\tilde{H})$. The odds ratio ϕ_F of course is the ratio of probabilities assessed by the experts on the basis of F. What does the statistician expect about this ratio, a-priori? Remember that the statistician does not even know precisely what the information possibilities defining F are. That is why a direct assessment of ℓ_F in terms of the ratio $P(F|H)/P(F|\tilde{H})$ would be so difficult. Nonetheless, he/she would have some idea of *how informative* this shared expert information might be. Otherwise there would be no reason to consult an expert. This is the value of recognising the equivalence of $\rho\ell_F$ with the conceptually accessible odds ratio $P(H|F)/P(\tilde{H}|F)$.

For the same reasons as (19), we also find it worthwhile to define two similar transformations, ϕ_1 and ϕ_2 :

$$\rho\ell_{G_1} = \frac{P_1(H|G_1)}{P_1(\tilde{H}|G_1)} \equiv \phi_1 \ , \quad \text{and} \quad \rho\ell_{G_2} = \frac{P_2(H|G_2)}{P_2(\tilde{H}|G_2)} \equiv \phi_2 \ . \tag{20}$$

Now a useful factorisation of T_i can be achieved via an argument which requires discussion:

$$\begin{aligned} T_i \quad &\equiv \quad \frac{P_i(H|FG_i)}{P_i(\tilde{H}|FG_i)} \quad = \quad \frac{P_i(FG_i|H)P_S(H)}{P_i(FG_i|\tilde{H})P_S(\tilde{H})} \\ &= \quad \frac{P_i(F|H)P_i(G_i|H)P_S(H)}{P_i(F|\tilde{H})P_i(G_i|\tilde{H})P_S(\tilde{H})} \\ &= \quad \rho \ \ell_F \ \ell_{G_i} \qquad\qquad \text{for each} \quad i = 1, 2 \quad . \end{aligned} \tag{21}$$

The first defining equality merely repeats the definition made in (6). The second equality, then follows from Bayes' theorem, using additionally the fact that S's asserted probability for H is the same as would be the assertions of experts 1 and 2, if they did not know for certain the precise expert information that they do, but rather merely B as S does. The third equality follows from an assertion of conditional idependence between F and each G_i given H and $\tilde{H}$ that we motivate in the paragraph that follows. The fourth and final equality of (21) makes use of the concise notation specified in (4), (5), (6) and (7). We now need digress on the motivation.

The truth or falsity of H rests within the context of conditions that convey confusing signals regarding its truth. If this were not the case, virtually every observant person who thinks about H and observes its context would be certain of its truth or falsity! Now the truth of F is a signal about H that is received and is shared by both of our experts, typically in the form of the likelihoods $\mathcal{L}(H;F)$ and $\mathcal{L}(\tilde{H};F)$. Nonetheless, this signal is *not definitive* about the truth or falsity of H. This is why the experts are still uncertain about H. To the statistician, who is not even accustomed to thinking about the definition of F, the truth of $\tilde{F}$ would be another possible signal about H. But one thing the statistician would say is this: whatever H signals about its truth via the truth or falsity of F, the additional evidential content H would signal about itself via the truth of either expert's specific experience, G_i, would be the same no matter whether F or $\tilde{F}$ were true. For S has no way of knowing whether the evidence that F and G_i provide about H are in support of its truth or put in doubt its truth, or are either one in support and the other leading to doubt. It is the truth or falsity of H itself that provides the basis for the information signal. This is the general motivation for the presumed conditional independence of F and each G_i.

As an example, consider again the incidence of Varroa mites in a sample from a beehive. H is the event that a sample of bees shows evidence of Varroa in a powdered sugar test today. F is the event that the sex ratio at birth of male to female honey bees in a healthy hive in the spring is around 1/25, a fact known to both bee experts. G_1 is the event known to expert 1 that this hive was treated with the insecticide "Capstan" last year. Whether or not G_1 is the case would be evidenced in a naturally confusing way by the truth of H. It is easy to imagine assessing a conditional probability such as $P(G_1|H)$. Whether or not F is *also* the case would be irrelevant to an assertion for the probability of G_1 *given the truth of* H. The assertion value of $P(G_1|FH)$ would be the same as that of $P(G_1|H)$. The same could be said given the falsity of H. This is the substantive content of the conditional independence in this context. Let us now return to our development.

Using equation (21), it is also clear that each $\rho T_i = \phi_F \phi_i$, and thus, $T_1 T_2 = \phi_F^2 \phi_1 \phi_2 / \rho^2$. Equivalently,

$$log(T_1 T_2) = 2log(\phi_F) + log(\phi_1) + log(\phi_2) - 2log(\rho) \quad . \tag{22}$$

Consideration of these three ϕ quantities together will allow us to specify the mixing function $F(\phi_F|(T_1 = t_1)(T_2 = t_2))$ that we desire; for $log(T_1T_2)$ is a linear function of $log(\phi_F), log(\phi_1)$ and $log(\phi_2)$.

For this and other reasons that will become clear, it is natural to consider the *lognormal* family of distributions to represent prior opinions *for the odds ratios* defined by the ϕ's. On account of equation (22), this form would allow a joint lognormal distribution for the ϕ's and the product T_1T_2 as well. As shall be seen, this would imply that the distribution for the probabilities $P(H|F)$ and $P(H|G_i)$ is joint *logitnormal*. It will be best to discuss these distributions in stages.

6.1 Specifying a joint lognormal distribution for $\phi_F, \phi_1, \phi_2, T_1$ and T_2

To begin, we shall merely specify the form of a joint distribution that requires the elicitation of three understandable parameters, σ_F^2, σ_G^2 and $c_{1,2}$. Then we shall discuss what the assessment of the statistician's opinions in this form would mean.

$$\begin{pmatrix} \phi_F \\ \phi_1 \\ \phi_2 \\ --- \\ T_1 \\ T_2 \end{pmatrix} \sim \; Lognormal \; \left(\begin{pmatrix} \mu(\pi,\, \upsilon_F) \\ \mu(\pi,\, \upsilon_G) \\ \mu(\pi,\, \upsilon_G) \\ ----------- \\ \mu(\pi,\, \upsilon_F) + \mu(\pi,\, \upsilon_G) - log(\rho) \\ \mu(\pi,\, \upsilon_F) + \mu(\pi,\, \upsilon_G) - log(\rho) \end{pmatrix} ,\; \Sigma \; \right) ,$$

where Σ =

$$\left(\begin{array}{ccc|cc} \sigma_F^2(\pi,\, \upsilon_F) & c_{i,F} & c_{i,F} & \sigma_F^2 + c_{i,F} & \sigma_F^2 + c_{i,F} \\ c_{i,F} & \sigma_G^2(\pi,\, \upsilon_G) & c_{1,2} & \sigma_G^2 + c_{i,F} & \sigma_G^2 + c_{i,F} \\ c_{i,F} & c_{1,2} & \sigma_G^2(\pi,\, \upsilon_G) & \sigma_G^2 + c_{i,F} & \sigma_G^2 + c_{i,F} \\ ----- & ----- & ----- & ------- & ------- \\ \sigma_F^2 + c_{i,F} & \sigma_G^2 + c_{i,F} & \sigma_G^2 + c_{i,F} & \sigma_F^2 + \sigma_G^2 + 2c_{i,F} & \sigma_F^2 + 2c_{i,F} + c_{1,2} \\ \sigma_F^2 + c_{i,F} & \sigma_G^2 + c_{i,F} & \sigma_G^2 + c_{i,F} & \sigma_F^2 + 2c_{i,F} + c_{1,2} & \sigma_F^2 + \sigma_G^2 + 2c_{i,F} \end{array} \right) . \tag{23}$$

The multivariate distribution (23) has been partitioned only to distinguish that the statistician's uncertainty about the first three component variables is to be assessed, while the cohering distribution with the fourth and fifth components is implied from the first three on account of the fact that each $T_i = \phi_f \phi_i / \rho$, and thus each $log(T_i) = log(\phi_F) + log(\phi_i) - log(\rho)$. The new parameters υ_F and υ_G and their roles in determining $\mu(\pi,\, \upsilon_F)$, $\mu(\pi,\, \upsilon_G)$, $\sigma_F^2(\pi,\, \upsilon_F)$ and $\sigma_G^2(\pi,\, \upsilon_G)$ that appear in this specification shall now be discussed.

We need to consider the statistician's initial opinion about the conditional odds ratios ϕ_F, ϕ_1 and ϕ_2 coherently in the context of the assertion $P_S(H) = \pi_S$ that has already been specified . Coherency requires that $P_S(H) = P_S[P_i(H|F)] = P_S[P_i(H|G_i)]$ for reasons we discussed around equation (8) at the end of section 4.

In motivating (23), consider marginally a distribution for the experts' shared odds ratio, ϕ_F. How *precise* will the common information F be regarding H? ... and how secure can a person be in *locating* its strength, knowing only the background information, B? The statistician's answer to these questions will be represented by a parametric choice of σ_F^2 which appears in equation (23). Although the form of that distribution for $\phi_F \equiv P_i(H|F)/P_i(\tilde{H}|F)$ is specified as lognormal, it will be easier to assess through direct consideration of possible values for the experts' assertion, $P_i(H|F)$. Under this transformation, the statistician's distribution for $P_i(H|F)$ is *logitnormal.*

6.2 A technical review of the logitnormal distribution

To think clearly about the lognormal distribution specified in (23) and to use it in practice, we need to recall a few relevant aspects of a related transformation, the logitnormal distribution. While the logit transformation of a variable on the unit-interval is well-known and has been studied in myriad applications, the Logitnormal distribution has escaped direct exposition. This is somewhat surprising given the widespread use of logistic regression analysis. The logitnormal distribution was suggested for statistical work in the seminal work of Lindley (1964) on contingency tables. It has subsequently been used in hierarchical Bayesian modelling as early as Leonard (1972) and in applications such as that of Kass and Steffey (1989). In all these cases it was applied in the form of the normal distribution for the logit transformation of a variable in the unit-interval, rather than directly in the form of a distribution over that interval.

If Θ is a variable within the unit interval, the following three specifications are equivalent:

$$\begin{aligned} \Theta &\sim Logitnormal(\mu, \sigma^2) \\ \Phi = \Theta/(1-\Theta) &\sim Lognormal(\mu, \sigma^2) \\ X = log[\Theta/(1-\Theta)] &\sim Normal(\mu, \sigma^2) \ . \end{aligned}$$

While the lognormal density for Φ is completely tractable, with moments specifiable in terms of μ and σ^2 (the moments of the normal density for X), the logitnormal moments of Θ require numerical evaluation. The logitnormal density for Θ derives easily from the normal density for X via the inverse transformation $\theta = \frac{e^x}{(e^x+1)}$, viz.,

$$f(\theta|\mu, \sigma^2) \;=\; \frac{1}{\sqrt{2\pi}\sigma\theta(1-\theta)} \; \exp\{-1/2[(\frac{log(\frac{\theta}{1-\theta}) - \mu}{\sigma})]^2\} \qquad \text{for } \theta \in (\,0,\ 1\,)\ . \qquad (24)$$

However the moments of this density cannot be integrated algebraically, and require numerical integration. We can study them graphically.

Figure 1 displays the mean and variance functions for the Logitnormal(μ, σ^2) family using the Normal(μ, σ^2) family's signal to noise parameter, μ/σ, and σ as arguments. Figure 2 displays contours of $(\mu/\sigma, \sigma)$ pairs that support constant values of the mean at $E(\Theta) = .6, .75$, and .9, and values of the variance at $V(\Theta) = .01, .04$, and .10. We have limited this display to non-negative values of μ/σ due to the obvious result that if $\Theta \sim Logitnormal(\mu, \sigma^2)$, then $(1-\Theta) \sim Logitnormal(-\mu, \sigma^2)$. The choice of μ/σ as the argument variable is for practical reasons in limiting the range required for the display.

The functions $E(\Theta) = M(\mu, \sigma^2)$ and $V(\Theta) = V(\mu, \sigma^2)$ are invertible, of course, so that a specification of $E(\Theta)$ would determine a contour of supporting (μ, σ^2) values from Figure 2, and a specification of $V(\Theta)$ would then identify the appropriate single (μ, σ^2) pair, selected from this contour. Once $E(\Theta)$ is specified, it is probably easier to specify $V(\Theta)$ by examining an array of densities that support this mean value, as we shall now see.

Figure 3 displays graphs of three logitnormal densities that support $E(\Theta) = .60$, to indicate what members of this family of densities can look like. It is interesting to see how they can be different from Beta family densities. When σ^2 is small, the densities are unimodal and do look similar to the Beta(α, β) densities when α and β both exceed 1. However the logitnormal densities always converge to 0 as $\theta \to 0^+$ and as $\theta \to 1^-$, and when σ^2 is large enough the densities are bimodal. In this way they differ markedly from the family of Beta(α, β) densities when α and β are both less than 1. These Beta densities have asymptotes at 0^+ and at 1^-. The logitnormal family contains no members that resemble the Beta densities when $\alpha \leq 1$ and $\beta \geq 1$, or vice versa.

We end this review by noting that the quantities represented by Θ in this technical review would be those designated θ_F, θ_1 and θ_2 in our experts' problem. That is, they are conditional probabilities tendered by the experts that are unknown to the statistician. We are imagining the statistician to assess them with a logitnormal distribution as expressed in (23). In this context it is worth recognising specifically that *large values of* σ^2 in the lognormal specification for the distribution characterise opinions that expect Θ to be *more informative*,

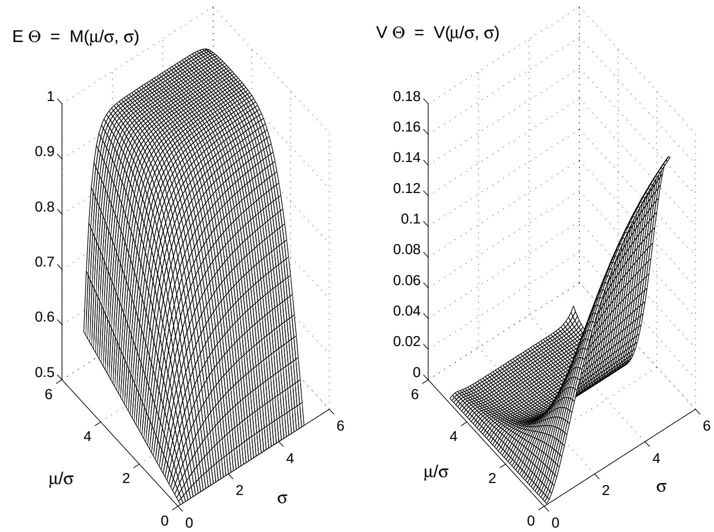

Figure 1: The graph at left depicts values of $E(\Theta)$ as a function of the Logitnormal μ/σ and σ parameters. At right appears the Variance of Θ as a function of these same parameters.

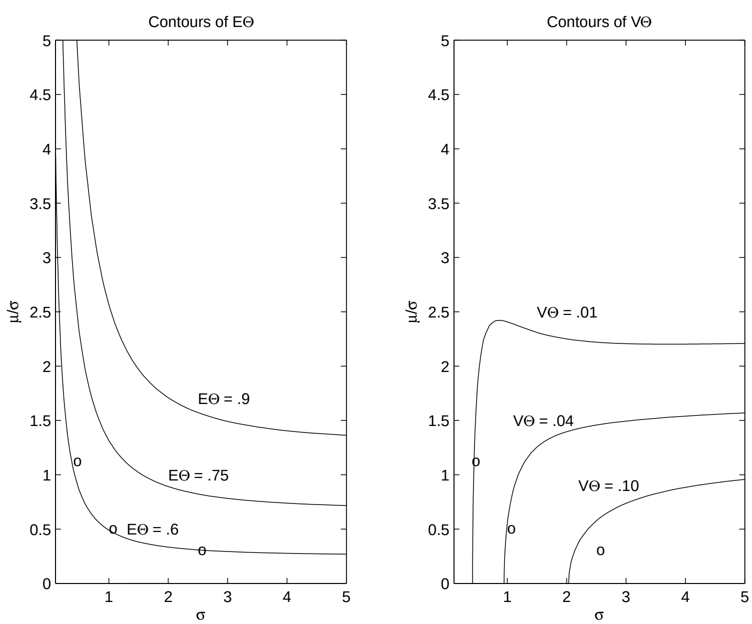

Figure 2: The graph at left shows contours of $(\mu/\sigma, \sigma)$ combinations that support constant values of $E(\Theta)$ at .6, .75, and .9. At right appear contours of constant values of the variance, $V(\Theta)$, at .01, .04 and .10. The logitnormal densities pertaining to the three parametric points exhibited by circles on the contour of $E(\Theta) = .6$ are displayed in Figure 3.

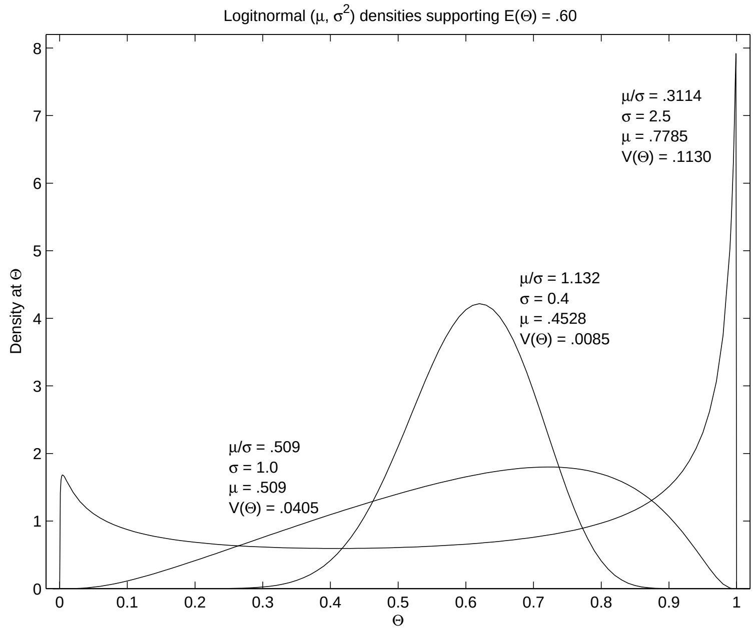

Figure 3: Logitnormal (μ, σ^2) densities with expectation $E(\Theta) = .60$ are specified by various (μ, σ) pairs. Again, higher values of σ specify a bimodal logitnormal density. The $(\mu/\sigma, \sigma)$ values that specify these three densities are indicated by small circles in Figures 1 and 2.

not less informative. For example, in Figure 3 of the densities with mean $E(\Theta) = .6$ the density for Θ bunches in intervals close to 0 and in intervals close to 1 when $\sigma = 2.5$, while it is more concentrated near .6 when $\sigma = 0.4$. This is not really counterintuitive, but it is a bit tricky to get straight in your mind. Think about it. When σ^2 is large, Θ is expected to be very informative, for it is expected to be a number (the probability for H assessed by the experts) that is close to either 0 or 1. But when σ^2 is small, the density for Θ is sharply bunched around .6, which is to say that the probability to be asserted by the experts is expected to be only mildly informative.

6.3 Motivating the parameter selection for the joint lognormal

Based on our knowledge of properties of the logitnormal distribution, we need to motivate a sensible specification of the parameters υ_F, υ_G, and $c_{1,2}$ that identify the multivariate logitnormal density of (23). The quantity in our problem that plays the role of Θ in the technical review is the experts' shared assertion value of $P_i(H|F)$. Having already assessed an expectation for this quantity via the initial assertion of $P_S(H) = \pi_S$, the statistician need now only choose a value for $V[P_i(H|F)]$ to have specified a distribution in the form of (23), which is lognormal for the odds ratio ϕ_F, or equivalently logitnormal for the probability $P_i(H|F)$. We denote the selected value of $V[P_i(H|F)]$ by υ_F in the covariance matrix of equation (23.) The choice of υ_F can be made most easily by examining densities that support this expectation, π_S, in the form of the contour of $M(\mu, \sigma^2) = \pi_S$ values. This is exemplified in the graph of logitnormal densities for which $E(\theta) = .60$. In this case, each of the densities with this mean is characterised by a different choice of both μ and σ. Examine Figure 3. A choice of υ_F, and thus of σ_F^2, can be made from the density that appropriately represents the statistician's uncertainty by visual inspection. We shall compute three examples based upon different expectations about expert information. The first assesses strong information in F, with $\upsilon_F = .1130$. The associated density is bimodal with peaks near 0 and 1, and a central relatively uniform trough. The second assesses moderate information in F, with $\upsilon_F = .0405$. It is unimodal but rather diffuse over a large region. The third will assess rather weak information in F, with $\upsilon_F = .0085$. The density appears roughly triangular over the interval (.4, .8), peaking around .6. In terms of the lognormal variance parameter, these translate to $\sigma_F = 2.5$ (strong), $\sigma_F = 1.0$ (moderate) and $\sigma_F = .4$ (weak).

The next feature of (23) the statistician needs to consider is his/her opinion about the relative strength of either $P_i(H|G_i)$ to the common $P_i(H|F)$. Remember that both of these conditional probability assertions have been assessed with the same expectation, $P_S[P_i(H|F)] = P_S[P_i(H|G_i)] = P_S(H) = \pi_S$. The relative sizes of $V[P_i(H|F)] \equiv \upsilon_F$ and $V[P_i(H|G_i)] \equiv \upsilon_G$ then depend on whether the statistician thinks the common information available to both experts is more or less informative about H than the specific information they hold separately. The resolution of this consideration determines υ_G and thus the value of σ_G^2. Again, the assessment is aided by visual inspection of the family of densities that support the same expectation value. In the solution to the problem presented here, we suppose that the same variance applies to our consideration of the information sources specific to each expert, so the joint distribution for ϕ_1 and ϕ_2 is an exchangeable one. Of course this is not required by coherency. It is easy to imagine problems in which one expert is considered to be more knowledgeable than another, though both are considered to have information worth exploiting. In such cases, distinct variance specifications, υ_{G_1} and υ_{G_2}, would be required.

Our third specification of $c_{1,2}$ requires an assessment of the mutuality of the *direction* of the evidence about H that is held specifically by the two experts, when specified in terms of their log-odds ratios. The imagined communality of the directions of their private information specifies the covariance between the two of them, denoted in the joint lognormal parameterisation (23) by $c_{1,2}$. Does the statistician imagine that the information sources privately available to the experts would be mutually supportive (or non-supportive) about H, or perhaps would they be opposing? (the one supportive, the other non-supportive) ... or would they be assessed independently? The size of $c_{1,2}$ would typically be easier to assess in the form of a correlation, either between $log\,\phi_1$ and $log\,\phi_2$, or between $P_1(H|G_1)$ and $P_2(H|G_2)$. The numerical sizes of the correlations between these transforms are surprisingly close. Computational examples appear in a systematic Table in a technical note by Frederic and Lad (2003). The sign of either correlation could be positive or negative. Positivity would be appropriate in problems in which the experts rely on similar experiences which are different from each other only in their precise details. Negativity could be appropriate in cases when the consulted experts are recognised as having very different and even contrasting experiential bases for their assertions. We shall comment further on this matter in our concluding discussion, for there are technical limits on the sizes of negative correlations that can be considered in the context of the problem we are studying here. Moreover, in some scenarios the influence of the correlation is minimal. Nonetheless we shall present numerical examples involving negative as well as positive correlations.

We are finally left to consider a covariance between $log\,\phi_F$ and $log\,\phi_i$, denoted in (23) by $c_{1,2}$. For the particular solution of trusting and cooperative science we consider here, we presume that this is determined so that the correlations between these two quantities are equal to the correlation between the log-odds ratios corresponding to the two specific information sources. The equality of these two correlations would imply that the covariance denoted by $c_{i,F}$ must equal $c_{1,2}\,(\sigma_F/\sigma_G)$.

Having discussed motivations for the determinations of $\sigma_F^2, \sigma_G^2, c_{1,2}$ and $c_{i,F}$, we have thus completely specified the joint lognormal distribution identified in (23). It remains only to comment that the distribution of the fourth (partitioned) component of the multivariate lognormal distribution in (23) derives from standard methods for linear combinations of normal variables, on account of the log-linear equation (22).

7 Computing $P_S(H|(T_1 = t_1)(T_2 = t_2))$

We can now use the distribution (23) to complete the integration we specified in equation (18) which we repeat here using the notation of $\phi_F \equiv \rho\ell_F$.

$$P_S(H|(T_1 = t_1)(T_2 = t_2)) \;=\; \int_0^\infty \frac{t_1 t_2}{t_1 t_2 + \phi_F}\; dF(\phi_F|(T_1 = t_1)(T_2 = t_2)) \qquad (18)$$

We need only identify the conditional distribution of ϕ_F given $(T_1 = t_1)(T_2 = t_2)$ from the bivariate lognormal distribution for (ϕ_F, T_1T_2) embedded in the multivariate lognormal distribution specified in (23). Normal theory tells us that this conditional distribution is also lognormal, with the conditional moment parameters

$$E\ [\ log(\phi_F)\ |\ (T_1 = t_1)(T_2 = t_2)\ ]\ =\ \mu\,(\pi,\ \upsilon_F) +$$
$$[\ log\,(t_1) + log\,(t_2) - 2\,\mu\,(\pi,\ \upsilon_F) - 2\,\mu\,(\pi,\ \upsilon_G) + 2log\,(\rho)]\ [\frac{(\sigma_F^2 + c_{i,F})\ (\sigma_G^2 - c_{1,2})}{D}] \tag{25}$$

and

$$V\ [\ log(\phi_F)\ |\ (T_1 = t_1)(T_2 = t_2)\ ]\ =\ \sigma_F^2 - \frac{2(\sigma_F^2 + c_{i,F})^2\ (\sigma_G^2 - c_{1,2})}{D}\ , \tag{26}$$

where

$$D\ =\ (\sigma_G^2)^2 + 2\sigma_G^2\sigma_F^2 + 4\sigma_G^2 c_{i,F} - 2\sigma_F^2 c_{1,2} - 4c_{i,F}c_{1,2} - c_{1,2}^2\ .$$

A final aspect of this long development requires us to realise that it is the conditional expectation of $t_1t_2/(t_1t_2 + \phi_F)$ which is specified by (18) rather than the expectation of $log(\phi_F)$. Although this expectation cannot be expressed in a closed algebraic form, the integration can be computed numerically for any specific values of $t_1t_2, \pi, \sigma_F^2, \sigma_G^2$ and $c_{1,2}$. It is interesting that this expectation is a weighted modification of the standard "odds to probability transformation": $prob = odds/(odds + 1)$. The modified transform declines from 1 to 0 as ϕ_F increases from 0 without bound. We shall discuss its interpretation after we examine some numerical results in the next two Sections.

8 Contours of equivalent inferences

The inference the statistician makes from hearing the probability assertions $P_1(H) = \pi_1$ and $P_2(H) = \pi_2$ depends only on the product of the odds ratios that their probabilities imply. (Remember too that this is a special case of the more general result which requires that the multiplicand odds ratios be raised to powers.) It is clear from equation (18) that $P_S(H|(T_1 = t_1)(T_2 = t_2))$ is constant for all pairs of the experts' probabilities (π_1, π_2) that yield the same value of the product t_1t_2. This means that in the space of all possible probability assertion pairs, the posterior probability for H is constant on contours for which $\pi_2 = t_1t_2(1 - \pi_1)/[t_1t_2(1 - \pi_1) + \pi_1]$. Of course, in any specific application, the constant conditional probability value on each contour depends on the specification of $\pi, \sigma_F^2, \sigma_G^2$, $c_{1,2}$ and $c_{i,F}$ by the statistician.

It is well worth examining a few contour lines of the constant posterior probability values to understand the procedure we shall follow in exhibiting the consequences of various possible prior specifications. Figure 4 displays contours of constant products of the experts' odds at values of $t_1t_2 = .01, .10, .40, 1, 2.5, 10$ and 100. For now, let us only notice an interesting feature of our conclusion thus far. As long as $\pi_1 + \pi_2 = 1$, the product of the experts odds ratios also equals 1. These pairs are identified by the contour line showing $t_1t_2 = 1$. Thus, for examples, the statistician will make exactly the same inferences from hearing the experts announce (.1, .9) as from hearing (.5, .5). Precisely what the inference will be depends on prior expectations, which we shall examine in three examples. However the other contours of equal inferential probabilities are not linear. If the *average* of the experts' probabilities were the sufficient statistic (instead of the product) the contours would be lines for which $\pi_1 + \pi_2 = K$. It is quite evident both geometrically and algebraically in this problem that they are not. The average of the experts' probabilities becomes more and more a distortion of the inferential content of their expertise as the product of their odds ratios increase or decrease away from 1.

Before proceeding to some numerical examples, notice that the contours of equivalent inferences from the experts' assertions of (π_1, π_2) are identical to the contours specified by the family of externally Bayesian pooling operators. Reviewing the identification of these operators in equation (1) shows that the pooling operator yields the same probability K for all (π_1, π_2) pairs specified by the contour

$$\pi_2\ =\ \frac{(\frac{K}{1-K})^{1/(1-w)}\ (1 - \pi_1)^{w/(1-w)}}{(\frac{K}{1-K})^{1/(1-w)}\ (1 - \pi_1)^{w/(1-w)}\ +\ \pi_1^{w/(1-w)}}\ . \tag{27}$$

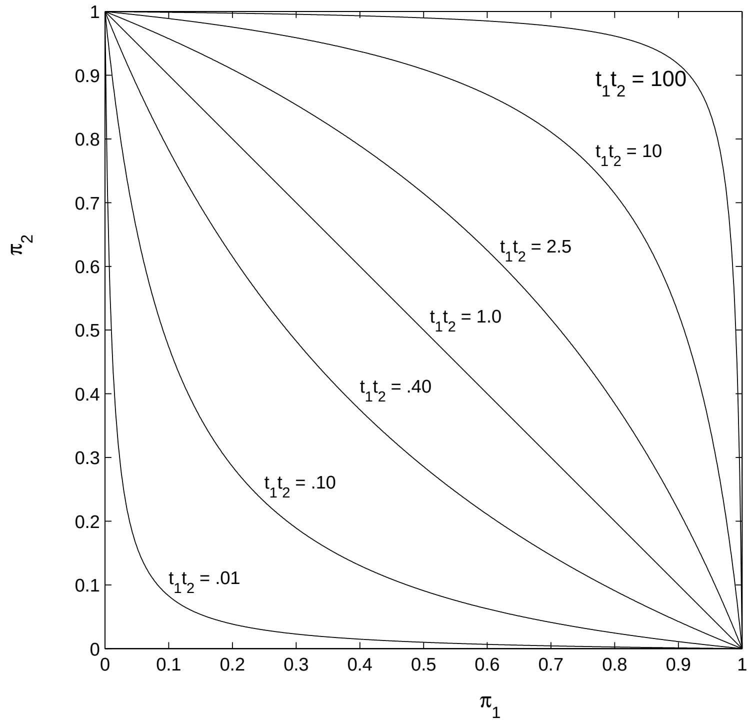

Figure 4: Contours of constant products of experts' odds ratios: $\pi_2 = \frac{t_1 t_2 (1-\pi_1)}{t_1 t_2 (1-\pi_1)+\pi_1}$

When $w = .5$, the form of this contour is identical with the contour determined by the product of the experts' odds ratios as a sufficient statistic. Thus, the coherent use of experts' probabilities we have characterised specifies contours of constant posterior probability for H that coincide with the contours prescribed by an externally Bayesian operator. However, we shall see shortly that the coherent function values on these contours are completely different from those associated with EB operators. (We also mention here without proof that the contours of the other EB operators with values of $w \neq .5$ coincide with those representing inferences when the (α, β) pairs specified in (13) that differ from $(1, 1)$.)

9 Numerical examples

Since the posterior probabilities are constant on contours through the (π_1, π_2) space, we need only to compute this probability once for each contour line. This is fortunate, because it will allow us to exhibit a lot of information from examples of inferences on a single graph. Since the sufficient statistic is the product of the odds ratios, $t_1 t_2$, we shall present graphs of inferential probabilities, $P(H|(T_1 = t_1)(T_2 = t_2))$ as functions of the square root of this product, transformed into a probability value which can vary from 0 to 1. Functionally, the argument of the inferential probability will thus be denoted by $\pi^* = \sqrt{t_1 t_2}/(\sqrt{t_1 t_2} + 1)$. *If both experts had asserted this same probability value*, the product of their odds ratios would be the same as was their actual product. We shall refer to this convenient computation as the experts' "equivalent shared probability," and denote it by $\pi^* \equiv \pi(t_1, t_2)$.

We need to stress here that the following analysis *does not presume* that both experts assert the same predictive probability. The same inference is made for any pair of experts' probabilities that lie on the same contour as does $(\pi_1, \pi_2) = (\pi^*, \pi^*)$. We defer discussion of the implications of the results for the principles of unanimity and compromise to the section that follows.

We now examine some numerical results of inference patterns using specific choices of assessments for the relative expected strengths of the shared information, F, and the individuals' specific information G_1 and G_2. The following three graphs are based on prior expectations of the strength of the shared and specific informations that are represented by the densities displayed in Figure 3. Each of these densities describes a situation in which the public information to the statistician mildly favours the occurrence of the event in question, for they all support $P_S(H) = .60$. Nonetheless, each density describes a different attitude toward the amount of information additionally available to the experts. The single peaked central density in that Figure (showing $V(\Theta) = .0085$) is the one that expects the information source to be the weakest; for the probabilities based on the information are expected to be only marginally above or below the expectation, which is .60. The bimodal density with the modes near the extreme values of 0 and 1 (showing $V(\Theta) = .1130$) is the one that expects the information source to be the strongest, for the probabilities based on the information are expected to be either closer to 0 or closer to 1, away from the expectation of .60. The milder unimodal density (showing $V(\Theta) = .0405$) represents a moderate level of expectation regarding the strength of the information, for it entertains substantial expectations of probabilities markedly above .60 and markedly below .60, likely to be as low as .20 and as high as .90. However, it highly supports probabilities near to .60 as well.

Each of the following graphs is based on a distribution in the form of equation (23) that picks a different pair of these densities to represent the statistician's expectation about the strength of the probabilities supported by the shared information, F, and the experts' specific informations, G_i. In each case, the probabilities based on the two experts' specific information are regarded exchangeably, with the correlation between them denoted by $\rho_{1,2}$. Moreover, the correlation between the probability supported by either specific expert information and the probability supported by the shared information is also assessed with this same correlation value. Thus, each graph will show the posterior probability functions $P[H|(T_1 = t_1)(T_2 = t_2)]$ based on the variances $V(\Theta_F)$ and $V(\Theta_{G_i})$ specific to that Figure, and with an array of specified correlation values covering $-.25, 0, .25, .5$ and .75.

9.1 Strong shared information, F, and weak specific G_i

The first results we examine represent the situation that the shared information is expected to be very informative, while the specific information of each expert is expected to be rather weak. The posterior probability functions in Figure 5 are ordered by the size of correlations they specify between the information sources.

For low values of $t_1 t_2$ (and thus of π^*) the highest function corresponds to the assessment of $\rho = -.25$, and the functions are layered lower and lower for values of $\rho = 0, .25, .5$ and .75, respectively. The five functions meet and reverse this order in the vicinity of an equivalent shared probability of $\pi^* = .68$. At this point, all the posterior probabilities equal about .69. At higher or lower values of π^*, however, the posterior probabilities are never really much different from π^* itself. The largest discrepancies occur at $\pi^* = .25$. (This implies a product of experts' odds ratios of 1/9, making a contour of $t_1 t_2 \approx .11$ in Figure 4.) In this case, the inferential probabilities based on the product of the experts' odds ratios equal to 1/9 are equal to 0.2736, 0.2579, 0.2444, 0.2323 and 0.2214, based on correlations of $-.25, 0, .25, .5$ and .75 respectively. When the product of the experts odds ratios equals $(.6/.4)^2$, whether or not it is based upon $\pi_1 = \pi_2 = .6$, the posterior inference of the statistician is very close to .6, though not precisely equal. The exact values are computed as 0.6131, 0.6080, 0.6042, 0.6012 and 0.5990, respectively, again based on the assessed value of $\rho_{1,2}$. Above values of $\pi^* = .68$ the functions reverse their order from smallest to largest, but they remain quite close to an inference agreeing with π^*.

It is worth remarking right here that all of these inference functions differ from the function that would correspond to an EB operator on π_1 and π_2. As can be seen from equation (1), that function would specify an exact 45-degree line: $g(\pi^*, \pi^*) = \pi^*$.

9.2 Moderate shared information, F, and weak specific G_i

The next results represent the situation that the shared information is still expected to be fairly informative, but less so than in Figure 5. Again, the specific information of each expert is expected to be weak, only mildly augmenting the basic information. To be precise, the inferences are based on the logitnormal densities with $E(\Theta) = .6$ and $V(\Theta_F) = .0405$ with $V(\Theta_{G_i}) = .0085$. Equivalently, $\sigma_F = 1.0$ while $\sigma_G = .4$. The posterior probability

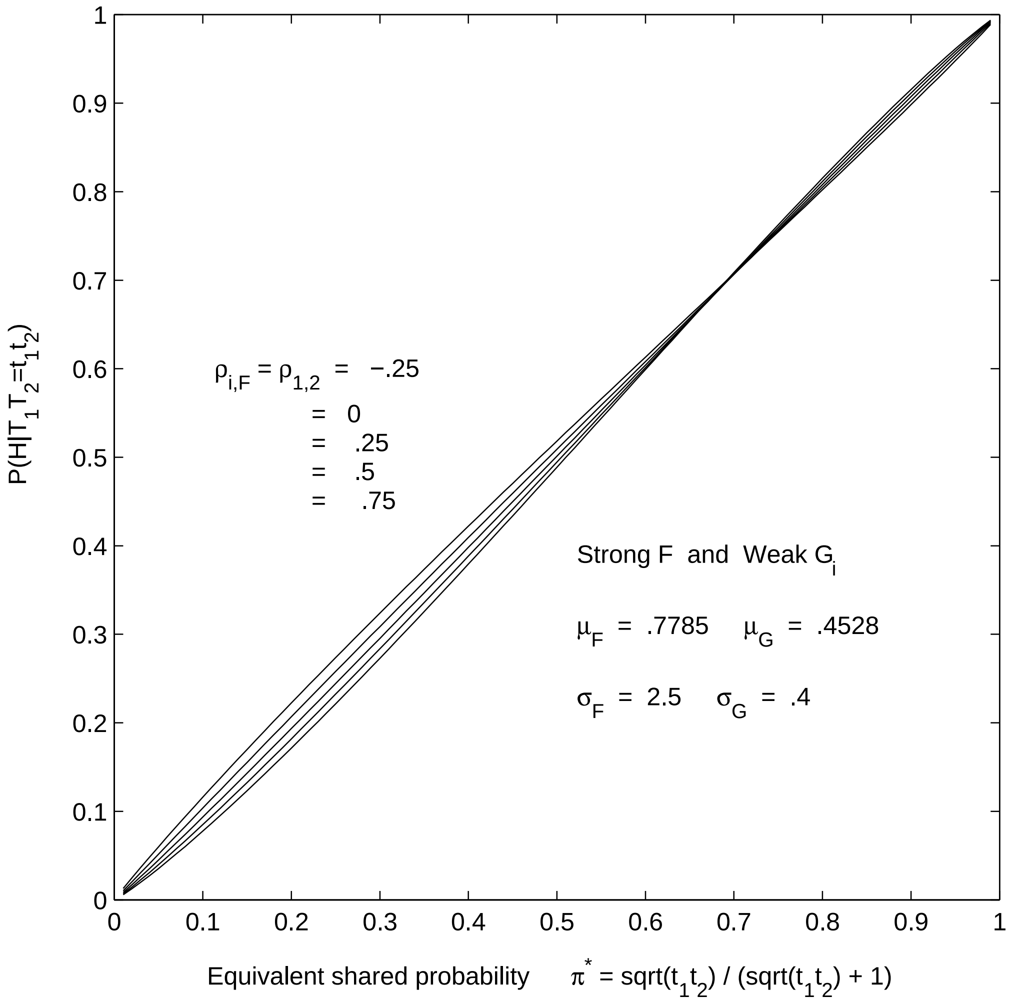

Figure 5: Posterior probability functions $P[H|(T_1 = t_1)(T_2 = t_2)]$ based on prior expectations of a strong information source in the shared information and weak information sources specific to the two experts.

functions displayed in Figure 6 are again ordered by the size of correlations they specify between the information sources. Again, for low values of $t_1 t_2$ (and its π^*) the highest function corresponds to the correlation value of $-.25$, and the functions are layered lower and lower for correlations of $0, .25, .5$ and $.75$, respectively. The functions are now a bit more separated from one another compared to Figure 5, and they now meet and reverse this order in the vicinity of $\pi^* = .64$. At this point, all the posterior probabilities also equal about .64. The largest discrepancies among the posterior probabilities based on different values of $\rho_{1,2}$ occur around $\pi^* = .25$. Here the inferential probabilities decline from .28 to .19 based on correlations of $-.25, 0, .25, .5$ and $.75$. Above values of $\pi^* = .64$ the functions reverse their order from smallest to largest, but they remain quite close to an inference agreeing with π^*.

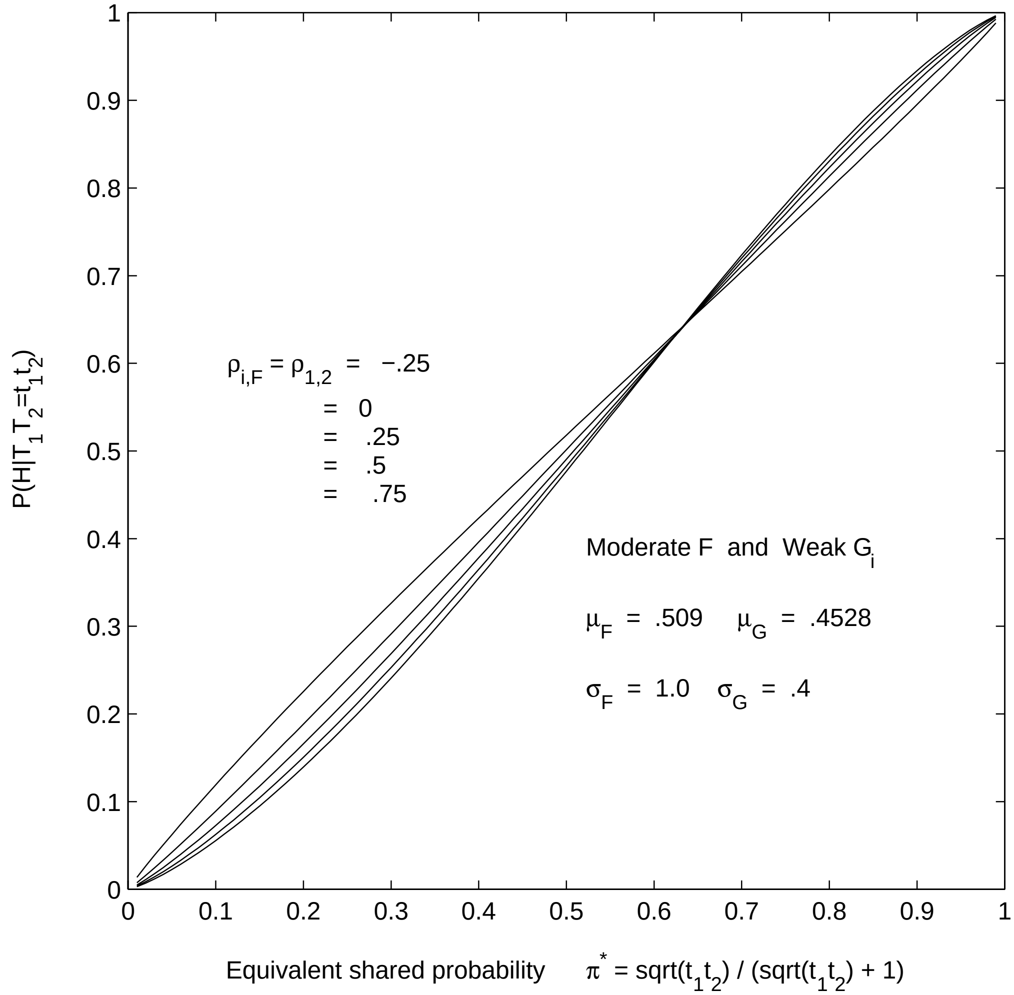

Figure 6: Posterior probability functions $P[H|(T_1 = t_1)(T_2 = t_2)]$ based on prior expectations of moderate shared information while still weak information sources specific to the two experts.

9.3 Weak shared information, F, and moderately informative G_i

The final results we display here are based on a scenario of expectation that only a little information is shared by the experts, while their specific information sources are regarded as being more promising, though not definitive. Figure 7 appears markedly different from the two figures we have already examined. The five functions based on different correlation values are now virtually indistinguishable to the eye. Moreover, rather than straddling the 45-degree line as in the previous two cases, the functions now ride well below that line before crossing it near $\pi^* = .6$. The five inferential probability functions are almost always equal to two decimal places. It is numerically insignificant but interesting that the order of the functions now changes, with the highest correlation-based function ($\rho_{1,2} = .75$) now being the highest function for low values of π^*, and being the lowest at high values. However the functions are no longer completely ordered by the sizes of the correlation either.

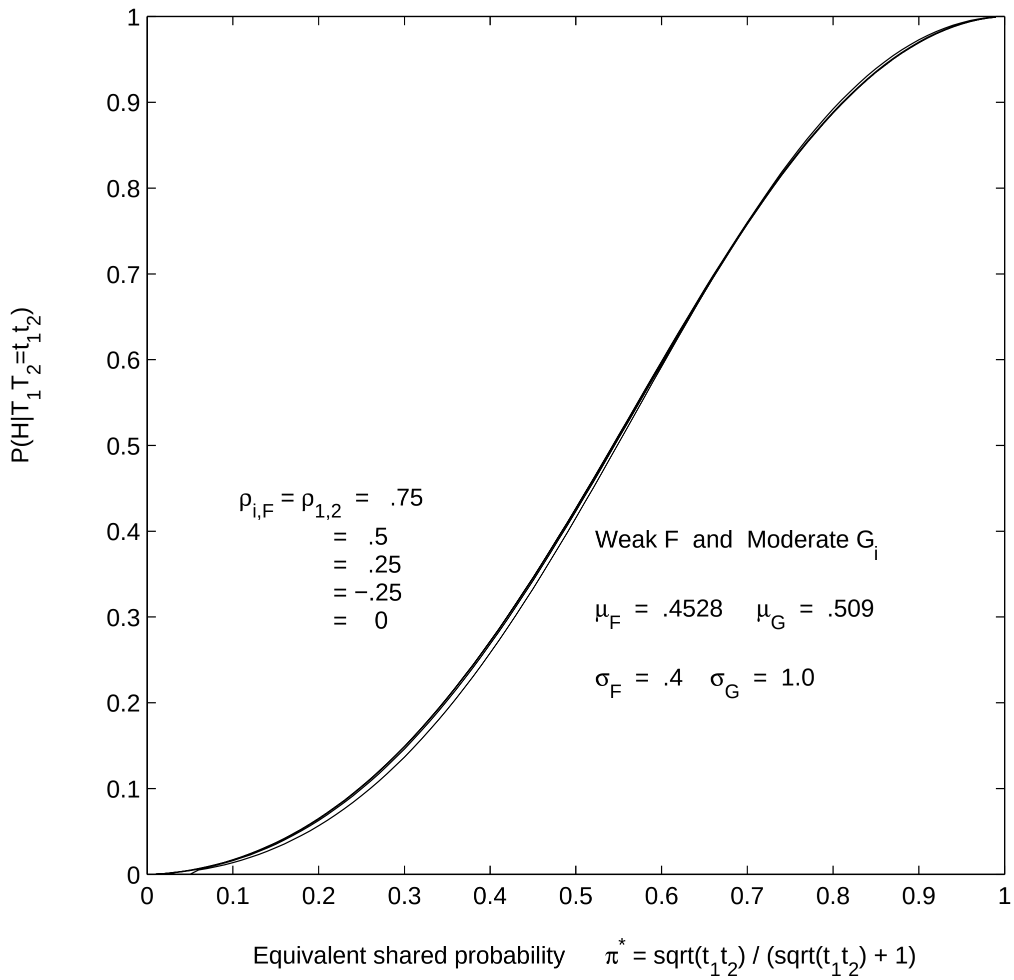

Figure 7: Posterior probability functions $P[H|(T_1 = t_1)(T_2 = t_2)]$ based on prior expectations of a weak shared information source while stronger information sources (though still moderate) are expected specific to the two experts.

9.4 An interpretation

The three graphical examples we have displayed, which are somewhat surprising, can be understood by thinking about the meaning of their parametric specifications. The interpretation does not arise in a straightforward manner from the algebraic specification of the result on account of the transformation from the logitnormal to lognormal specification and the numerical integration required to achieve the functional solution. Nonetheless, some motivation for the interpretation can be found directly in the algebraic schema as well. We shall firstly discuss the interpretations via the meaning of the specifications, and then remark on some algebraic aspects without belabouring every gruesome detail.

The stipulation that σ_F^2 is much larger than σ_G^2 signifies that the information common to both experts is considered to be much greater than that specific to each. *The common information motivates both* of their probability assertions, *but it should be used only once* as a basis for the statistician's inference rather than twice. Thus, since the additional information available to the experts individually is weak, the statistician's inferred probability for H given the product-odds ratio, should not differ very much from the experts' equivalent agreed probability, π^*. This is observed in both Figures 5 and 6. Both of these graphs show the conditional probabilities enveloping the line $P(H|T_1T_2 = t_1t_2) = \pi^*(t_1, t_2)$ for various values of the shared correlation $\rho_{1,2} = \rho_{i,F}$. When it is further specified that σ_F^2 is large itself, rather than merely large in relation to σ_G^2, the strength of the common information, even though counted only once, is such that it is only mildly affected by the information in G_1 and G_2. This is why the envelope displayed in Figure 5 is somewhat tighter than that in Figure 6. The ordering of the functions forming the envelope is also understandable. In the context of greater information expected in common, the marginal addition of information from G_1 and from G_2 would be expected to cancel each other when the correlation $\rho_{1,2}$ is negative, while the two sources are expected to accentuate each other when the correlation is positive. Thus, when the product T_1T_2 is observed to be less than expected, adjustments of the conditional probability downward from π^* increase as the specification of the correlation increases. This is also observed in both Figures 5 and 6.

When the strength of each expert's separate information source is expected to exceed that of their common source, the situation is rather different. On the one hand there is less information in each of the experts' assertions that needs to be diminished for the inference on account of "double counting" the shared information. Thus, the two assertions contain relatively more information than when they share substantial motivating information sources. This difference can be observed in comparing Figures 6 and 7. In Figure 7, the two experts' distinct assertions being different from their expected value of π_S conveys relatively more information than when the experts share greater common information, in Figure 6. Thus, the adjustments of the the conditional probability away from the "equivalent shared probability" is greater in this case. Differences among inferences based on the range of correlations specified are much smaller, as we observe in Figure 7. Nonetheless, it is sensible that the less correlated are these different information sources expected to be, the greater should be the adjustment away from the equivalent shared probability.

An algebraic analysis of the lognormal conditional distribution of $\phi_F|(T_1T_2 = t_1t_2)$, identified in equations (25) and (26), supports this interpretation. Without displaying details, we report here only that the adjustments to the conditional mean occur through a regression coefficient that is an increasing rational function of the variance ratio σ_F^2/σ_G^2. Analysis across the spectrum of ratio values supports the interpretation we have explained. Interestingly, it also shows that the sign of the partial derivative of the coefficient with respect to the correlation $\rho_{1,2}$ changes from negative to positive as the variance ratio declines through 1/2 toward 0. Exact detail of the interpretation is modified by corresponding changes in the conditional variance of ϕ_F and is somewhat obfuscated by the log and logit transformations and the numerical itegration required.

10 The status of unanimity and compromise

Figures 5-7 show that the principle of unanimity does not generally hold for inference from the two experts' predictive probabilities. Otherwise the graphs of $P(H|T_1T_2 = t_1t_2)$ as functions of π^* would be the 45° line in every case. Whenever $P(H|T_1T_2 = t_1t_2) \neq \pi^*$, the unanimity principle does not hold. Moreover, a graphical exposition can display precisely the extent to which the principle of compromise does and does not hold, as well.

Suppose that $P(H|T_1T_2 = t_1t_2) = \pi_G < \pi^*$. Figure 8 depicts such a situation when $\pi_G = .25$ and $\pi^* = .4$. Any (π_1, π_2) pair lying on that section of the equivalent inference contour containing points northeast of (π_G, π_G) exemplifies a pair of experts' assertions for which the principle of compromise does not hold. A result they exemplify is general. When coherent inference delineated by $P(H|T_1T_2 = t_1t_2) = \pi_G < \pi^*$ does not subscribe to unanimity, then neither will the principle of compromise pertain to any pair of probabilities on the equivalent inference contour for which which $min(\pi_1, \pi_2) \in [\pi_G, \pi^*]$. When $\pi_G > \pi^*$, a similar statement holds with $min(\pi_1, \pi_2) \in [\pi_G, \pi^*]$ replaced by $max(\pi_1, \pi_2) \in [\pi^*, \pi_G]$.

11 Concluding discussion

We shall conclude with one technical remark and three brief substantive comments. The results seem to provide much scope for discussion and further development.

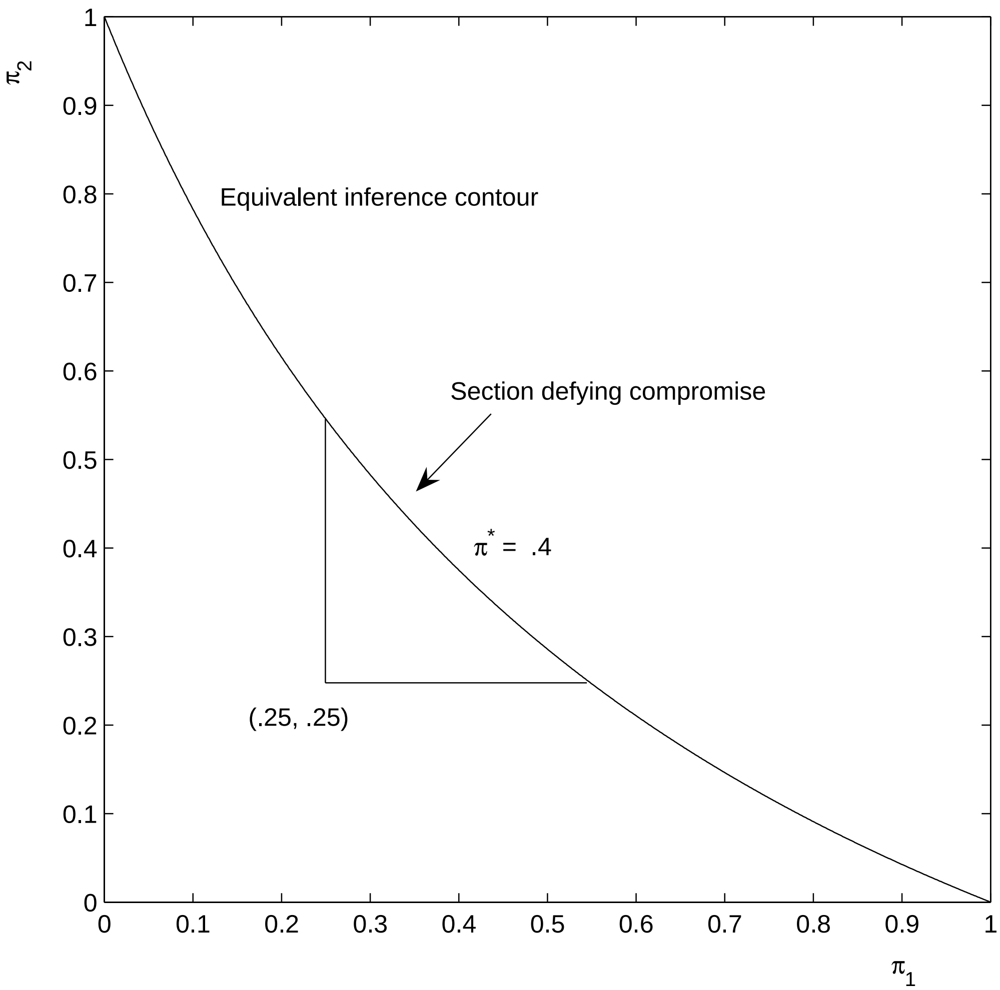

Figure 8: The arc of experts' probability pairs (π_1, π_2) that do not support the compromise principle in the condition that $P(H|T_1T_2 = t_1t_2) = \pi_G = .25 < \pi^* = .4$ contains those pairs for which $min(\pi_1, \pi_2) \in [\pi_G, \pi^*]$.

Technically, it is worth mentioning that there is a limitation in the degree of correlations that can be specified between the content of the expert-specific information sources and between either of these sources and the shared information, $\rho_{1,2}$ and $\rho_{i,F}$. It is evident from equation (23) that the likelihood ratios ϕ_1 and ϕ_2 are regarded exchangeably. Examination of (23) allows that although ϕ_F is not exchangeable with the ϕ_i in that distribution, the linear transformation $(\sigma_G/\sigma_F)[\phi_F - \mu(\pi, \nu_F) + \mu(\pi, \nu_G)]$ is exchangeable with the ϕ_i. For this reason, both of the two correlations must exceed $-1/2$. For technical explanation, see Lad (1996, p. 387). This is why the minimum specified common value of $\rho_{1,2}$ that we display in Figures 2, 3, and 4 was $-.25$.

Substantively, we begin with a remark regarding common allusions to a second opinion as an "independent opinion." The details of our analysis seem to clarify what can and should be meant by such a characteristic of the second opinion. It does not mean that our expectation of the desired second opinion is assessed as stochastically independent of the first. There are many many reasons why this would virtually never be the case. In none of the scenarios that we have displayed here numerically has stochastic independence been involved. However what would be desirable about a second opinion is that it be motivated by an information base that differs as much as possible from the base that motivates the first opinion. This is the feature that distinguishes the truly informative second opinion used for inference in Figure 7 from the second opinions underlying Figures 5 and 6.

We must comment specifically on the result that coherent inference does not require the combination of experts' opinions via an externally Bayesian operator, nor need it even support the principle of unanimity. The reason for the coherent defiance of EB operators is that when an expert asserts a probability value without an extensive explication of all evidence-based motives, the user of this probability does not know which nor how much of the evidence arises from each of the three types of sources we have characterised as B, F and G. Thus, a clear posterior inference based on further specific data can make use of it in a different way than it must when this same data content is known to inform already the experts who have been consulted. It is the disentangling of information shared by the experts from the information specific to each expert that is the real problem of making inference from experts' assertions. Our derived result achieves a resolution to this problem in an interpretable formal way.

A final remark is in order regarding a concept that we believe has provided a red herring in a sizeable literature on the combination of probabilities. We refer to the questionable desireability of the so-called "calibration" of assessors over an arbitrary range of probability assertions, an issue that has a fairly long history. It should be noted explicitly that none of the analysis we have presented has anything at all to do with the so-called recalibration of experts' probabilities. Indeed, we believe that the popularity of this concept derives from incomplete analysis of the issue. The review of French and Rios-Insua (2000, 4.21-4.23, pp. 122-125) discusses literature on the topic. However, it fails to address and even to reference the critical work on the calibration of probabilities that has been aired in Lad (1984, 1996 Section 6.6), Hill (1984) and Blattenberger and Lad (1985). Properly understood, all coherent probabilities of any assessor are well-calibrated. The arguments that challenge the widely referenced concept of calibration of probabilities have not really been addressed in the statistical community beyond the offhand dismissal by Shafer (1999, p. 648). We believe they deserve more serious consideration.

Acknowledgements

This research was begun while Frank Lad was a visiting professor of statistics at the Institute for Advanced Study, University of Bologna, during May and June, 2002. Special thanks to four reviewers of this article who helped us place our result in the context of published work and to simplify their generation. Thanks also to Dennis Lindley and Bob Winkler for helpful correspondence and to John Fountain for helpful discussion.

REFERENCES

Blattenberger, G. and Lad, F. (1985) Separating the Brier Score into calibration and refinement components: a graphical exposition, *The American Statistician*, **39**, 26-32. Correction **39**, 158.

Clemen, R. (1987) Combining overlapping information, *Management Science*, **33**, 373-380.

Clemen, R. and Winkler, R. (1990) Unanimity and compromise among probability forecasters, *Management Science*, **36**, 767-779.

Clemen, R. and Winkler, R. (1999) Combining probability distributions from experts in risk analysis, *Risk Analysis*, **19**, 187-203.

Dawid, A.P., DeGroot, M. and Morterra, J. (1995) Coherent combination of expert opinion (with discussion), *Test*, **4**, 263-313.

Dickey, J.M. (1980) Beliefs about beliefs, a theory of stochastic assessments of subjective probabilities, in *Bayesian Statistics: proceedings of the first Valencia international meeting, 1979*, J.M. Bernardo et al. (eds.), Valencia: University Press, pp. 471-489.

Frederic, P. and Lad, F. (2003) A technical note on the logitnormal distribution, University of Canterbury Department of Mathematics and Statistics Research Report UCDMS 2003/7, www.math.canterbury.ac.nz, 1-8.

French, S. and Rios Insua, D. (2000) *Statistical Decision Theory*, London: Arnold.

Genest, C., McConway, K. and Schervish, M. (1986) Characterization of externally Bayesian pooling operators, *Ann. Stat.*, **145**, 487-501. Reprinted in Kadane, J., Schervish, M. and Seidenfeld, T. (eds.) *The Foundations of Statistics Reconsidered*, 2002, Cambridge: Cambridge University Press.

Good, I.J. (1979) On the combination of judgements concerning quantiles of a distribution with potential application to the estimation of mineral resources, *J. Statist. Comput. Simul.*, **9**, 77-79.

Hill, B. (1984) Discussion of article by DeGroot and Eriksson, *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*, pp. 30-31.

Kass, R. and D. Steffey (1989) Approximate Bayesian inference in conditionally independent hierarchical models (Parametric empirical Bayes models), *Journal of the American Statistical Association*, **84**, pp. 717-726.

Lad, F. (1996) *Operational Subjective Statistical Methods: a mathematical, philosophical, and historical introduction*, New York: John Wiley.

Lad, F. (1984) The calibration question, *B. J. Phil. Sc.*, **35**, 213-221.

Lad, F. and Di Bacco, M. (2002) Assessing the value of a second opinion: the role and structure of exchangeability, *Ann. Math. Art. Int.*, **35**, 2002, 227-252.

Leonard, T. (1972) Bayesian methods for binomial data, *Biometrika*, **59**, pp. 581-589.

Lindley, D.V. (1964) The Bayesian analysis of contingency tables, *Annals of Statistics*, **35**, 1622-1643.

Lindley, D.V. (1985) Reconciliation of discrete probability distributions, with discussion, *Bayesian Statistics 2: proceedings of the second Valencia international meeting, 1983*, J.M. Bernardo et al. (eds.) Amsterdam: North-Holland, pp. 375-390.

Lindley, D.V.,Tversky, A. and Brown, R.V. (1979) On the reconciliation of probability assessments (with Discussion), *J. R. Statist. Soc.*, A, **142**, 146-180.

Madansky, A. (1964) Externally Bayesian groups, *Rand Corporation Memo* Rm-4141-PR, Santa Monica: Rand.

McConway, K. (1981) Marginalisation and linear opinion pools, *J. Amer. Statist. Ass.*, **76**, 410-414.

Morris, P.A. (1974) Decision analysis expert use, *Management Science*, **20**, 1233-1244.

Shafer, G. (1999) Book review of Operational Subjective Statistical Methods, *J. Amer. Statist. Ass.*, **94**, 648-649.

Winkler, R. (1968) Consensus of subjective probability distributions, *Management Science*, **15**, 61-75.

Winkler, R. (1981) Combining probability distributions from dependent information sources, *Management Science*, **27**, 479-488.

Zeckhauser, R. (1971) Combining overlapping information, *Journal of the American Statistical Association*, **66**, 91-92.