We can now proceed with the proof of Theorem 1. Suppose the DM's prior for X before seeing the data $Y = 0$ from Lemma 1 is f_1. The value of the marginal density of Y at 0 is

$$\int_S l(x) f_1(x)\, dx = \int_S f_2(x)\, dx = 1.$$

Since the marginal density of Y at the observed value 0 is finite, the posterior is the ratio of the joint density to the marginal, and f_2 is the posterior. Let the f in (1) be f_1. Equating the third term in (1) to the last term of (1) gives

$$\phi[c_2 lg;\, f_1] = \phi[c_2 lg;\, f_2], \tag{11}$$

where g is the expert's prior. Let Z be data, independent of Y, whose likelihood is $1/l$. Assume that the DM and the expert agree on this likelihood and that only the expert sees the data. Update according to the second term in (1) on both sides of (11), to obtain $\phi[g; f_1] = \phi[g; f_2]$. This is the conclusion stated in Theorem 1. Q.E.D.

References

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

GENEST, C. AND SCHERVISH, M., "Modelling Expert Judgments for Bayesian Updating, Technical Report" 304, Department of Statistics, Carnegie-Mellon University, 1983.

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

MORRIS, P. A., "Combining Expert Judgments: A Bayesian Approach," *Management Sci.*, 23 (1977), 679–693.

———, "An Axiomatic Approach to Expert Resolution," *Management Sci.*, 29 (1983), 24–32.

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

MANAGEMENT SCIENCE
Vol. 32, No. 3, March 1986
Printed in U.S.A.

CALIBRATION AND THE AGGREGATION OF PROBABILITIES*

ROBERT T. CLEMEN
College of Business Administration, University of Oregon, Eugene, Oregon 97403-1208

In order to avoid the task of assessing a complicated likelihood function, Morris uses an axiomatic approach to develop a multiplicative rule for aggregating a decision maker's and an expert's probabilities. An essential shortcoming of the multiplicative rule is that it does not allow the decision maker to model his beliefs about the dependence between his assessment and the expert's. The root of the problem lies in the fact that the decision maker must calibrate the expert's information. When the calibration is done properly, the decision maker is forced to tackle the task which Morris proposes to avoid.
(CALIBRATION; EXPERT RESOLUTION; DEPENDENCE)

Suppose a decision maker (DM) is interested in the outcome of a random variable (x) and solicits information about x from an expert (E). One way for DM to combine E's information $g(x)$ with his own prior assessment $f(x)$ is to apply Bayes' rule, treating g as a piece of data and assessing a likelihood function for it (Morris 1974). The likelihood function, assessed conditionally upon f, allows DM to model his beliefs

*Accepted by Robert L. Winkler; received March 12, 1985.

0025-1909/86/3203/0312$01.25

about the dependence between his own information and E's. Morris (1983, p. 24) points out that this likelihood function can be "extremely complicated and virtually impossible to assess in all but the simplest cases." To get around this problem, he develops an axiomatic approach to the DM's problem of combining g and f with the result that DM can use a multiplicative rule to combine the probability distributions.

Morris's axiomatic approach, however, does not allow DM to model the dependence between f and g (see French 1980). To see this, consider a simple example in which $f(x) = kx^s(1-x)^{n-s}$ for $x \in (0, 1)$, with k a normalization constant. Suppose this prior has arisen from DM observing s successes in n trials from a Bernoulli process with proportion x and then combining this information with a uniform prior for x. Now suppose E reports $g(x) = kx^s(1-x)^{n-s}$. If E has observed trials from the same Bernoulli process and has arrived at g in the same way that DM arrived at f, then DM's posterior f^* clearly depends on what he believes about the overlap between their data sets. If they both observed exactly the same data, then $f^* = f$, and E's information is completely uninformative to DM. (Alternatively, $f^* = g$ and DM's opinion is uninformative once he hears from E.) On the other hand, if DM believes that their data sets are independent, then $f^* = kfg$, exactly the result given by Morris's multiplicative rule. Furthermore, if DM believes that their data sets have some intermediate degree of overlap, then f^* will be a mixture of beta distributions as shown in Clemen (1984).

This example clearly violates Morris's multiplicative rule which essentially says that f^* must be the normalized product kgf in any event. It is tempting to say simply that the problem is with Morris's Axiom A, which states that f^* "should not depend on who observes a given piece of data if there is agreement on the likelihood function" (Morris 1983, p. 25). Apparently it does matter who observes which data; in particular it matters to what extent E's and DM's observations overlap. However, I believe the problem is more subtle than this. Note that I could have given an example using normal distributions and accomplished the same result. Then I would have apparently violated not only the axiomatically-derived multiplicative rule but also the multiplicative rule for normal distributions developed on the basis of Bayes' rule in Morris (1977). This earlier version requires DM to calibrate E's information g and then combine it multiplicatively with f. Similarly, the axiomatic approach requires g to be calibrated. I believe the problem lies in exactly what calibration means when it is done subjectively by a knowledgeable DM.

Traditionally calibration of expert opinions has been taken to mean an empirical procedure by which expert probabilities are adjusted so that they agree with the expert's past assessment record. An expert so treated is said to be "frequency calibrated." For a review of this kind of calibration see Lichtenstein, Fischoff and Phillips (1982). Morris (1977) distinguishes between this procedure and subjective calibration. Subjective calibration is a delicate concept because any subjective procedure must be consistent with coherence principles. In particular it must be consistent with Bayes' rule in the case of combining information. If E's assessments are subjectively calibrated by DM, this means that DM would not further adjust g. If DM has an uninformative prior, this definition of calibration makes sense, and Morris (1983, p. 27) captures this idea in his Axiom C: "If the decision-maker has a uniform [uninformative] prior, he should adopt a calibrated expert's prior as his own."

However, if DM has prior information, what does calibration mean? It appears to me that DM might make two different kinds of calibrating adjustments. The first kind adjusts E's assessments to square with his assessment record or with DM's corresponding subjective beliefs. The second kind involves adjusting E's information to account for any dependence between f and g. For example, DM might adjust E's information to eliminate any "double counting" of commonly observed data as in our example above. The calibration function $C(x)$ discussed in Morris (1977) accomplishes both of these tasks at once, provided that $C(x)$ is assessed conditionally on f (which it must be to preserve consistency with Bayes' rule). Once this calibration is performed, it may indeed make sense to combine f and the calibrated g_c multiplicatively.

The entire procedure that DM performs can perhaps best be thought of, in Morris's terms, as "jointly calibrating" E's and DM's opinions; the output of this procedure would be a probability distribution, like Morris's composite prior for multiple experts (Morris 1977, equation (3)), that reflects DM's beliefs about how DM and E perform jointly when they each assess a probability distribution for x. Once the joint calibration is performed, DM can take the jointly calibrated distribution as his own because all of his information has been encoded in that distribution.

But now consider what has happened. No matter how the calibration is performed, DM has had to model carefully the dependence between DM's and E's information. The procedure must be equivalent to applying Bayes' rule, or it would not be coherent. In attempting to escape the assessment of the likelihood function we find that DM still has an equivalent task.

Finally, equipped with the ideas of joint calibration and dependence between f and g, we can explain Morris's seemingly paradoxical example that apparently led him to the idea that discrete and continuous cases must be treated differently. If DM believes $P(\text{rain}) = 0.55$ and E reports $P(\text{rain}) = 0.55$, DM's posterior probability depends on what he believes about the way his forecast interacts with E's. If he believes that his own information is completely subsumed and properly used by E, then it makes sense that his posterior probability of rain would be 0.55; DM's assessment adds nothing new to the information provided by E. On the other hand if DM believes that his forecast includes information beyond that used by E, then he may indeed have a posterior probability that is different from 0.55, depending on the probability model used for combining the information. In any event, DM has to assess his beliefs about the dependence between the two statements in order to combine them. Whether we call his task joint calibration or application of Bayes' rule is immaterial, as the task is essentially the same.

References

CLEMEN, R. T., "Combining Overlapping Information: A Bayesian Approach," Ph.D. Dissertation, Indiana University, 1984.

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

MORRIS, P. A., "Decision Analysis Expert Use," *Management Sci.*, 20, 9 (1974), 1233–1241.

———, "Combining Expert Judgments: A Bayesian Approach," *Management Sci.*, 23, 7 (1977), 679–693.

———, "An Axiomatic Approach to Expert Resolution," *Management Sci.*, 29, 1 (1983), 24–32.

LICHTENSTEIN, S., B. FISCHOFF AND L. PHILLIPS, "Calibration of Probabilities: The State of the Art to 1980," in *Judgment under Uncertainty: Heuristics and Biases*, D. Kahneman, P. Slovic and A. Tversky (Eds.), Cambridge University Press, 1982, 306–334.