

Expert Judgment Elicitation Methods
Stephen C. Hora, University of Hawaii at Hilo
Hora@Hawaii.edu

1. Overview.....	1
2. Research Accomplishments	2
2.2. Performance of Aggregation Rules	3
2.3. Calibration of Multivariate Densities	4
3. Applied Relevance.....	6
3.1. Bioterrorism.....	6
3.2. Aggregation of Expert Judgments	6
3.3. Modeling Antiterrorist Budget Allocation Decisions as a Two Person Game	6
4. Collaborative Projects.....	6
5. Research Products.....	7
5.1. Publications and Reports	7
5.2. Presentations.....	7
5.3. Tools.....	7

1. Overview

Expert judgments are often required in the quantification of risk and decision models related to terrorist threats. This is the case when data are inconclusive or there is controversy about how evidence should be interpreted. When possible, multiple judgments are obtained as the differences among expert opinions portray the uncertainty in our knowledge concerning the target quantity.

The work accomplished by this project spans several areas. These areas include:

- Development of methods for assessing multivariate probability distributions for split fractions
- Creation of the a methodology to evaluate the performance of various aggregation rules
- Extensions of calibration concepts to multivariate distributions
- Modeling terrorist/government allocation decisions as two person game with spreadsheet implementation

The work on split fractions has generated an elicitation protocol based on decomposition of an k-variate Dirichlet distribution into k univariate beta distributions. This decomposition can significantly reduce the complexity of an assessment in that assessing k univariate densities is most often much simpler than assessing one k-variate density. This work has also generated decompositions that group possible outcomes together so that a general tree structure can be used for the decomposition instead of a strictly binary tree. Another extension of the methodology is the extension to other families of distributions beyond the Dirichlet. This entails using a generic distribution on [0,1] to model conditional frequencies or probabilities on the event tree.

The second area entails a new approach to evaluating the performance of aggregation rules. This method is based on the tools of mathematical statistics and employs infinite sequences of density functions that have predefined properties. Multiple infinite sequences are used to evaluate the performance of the aggregation. The sequences are designed so that the calibration of the modeled experts is controlled in such a way that they can be well-calibrated, overconfident, underconfident, or the experts can exhibit varying degrees of calibration. Similarly, experts with varying degrees of expertise (information) can be modeled and the degree of dependence among the experts modeled. The impact of varying these properties has been assessed using arithmetic and geometric aggregation of uniform, normal, and exponential densities.

Calibration of expert densities refers to the faithfulness of the judgments as empirically verified. Judgments that give intervals of a predefined probability should contain the true value with that predefined probability. The concept of calibration has not previously been extended to multivariate probability densities. This work has done so with the goal of providing minimal necessary and sufficient conditions and methods for assessing the calibration of multivariate densities.

The fourth area of work does not deal with expert judgments, but instead entails probabilistic modeling of terrorist threats in a two actor game environment. This work will eventually lead to the need for quantification of threats in terms of probabilities of terrorist success when implementing a strike program against a particular target. A spreadsheet based version of the model is in a prototype stage and operates with rather simple probability models. Further refinement of the probability modeling is underway.

2. Research Accomplishments

2.1. Split Fractions

Split fractions treated as relative frequencies present a special challenge in probability elicitation. A protocol has been developed for eliciting information about split fractions based on the decomposition of an $m-1$ dimensional Dirichlet density into $m-1$ univariate beta densities. The decomposition results in rather simple assessment questions. The aggregation of the recomposed Dirichlet distributions across multiple experts is also evaluated and arithmetic and geometric averages compared.

Extensions of this work that are completed or partially completed include:

- Using alternative univariate densities for binary bifurcations in the split fraction tree
- General decompositions of multivariate densities where the variables are constrained to sum to one

Figure 1

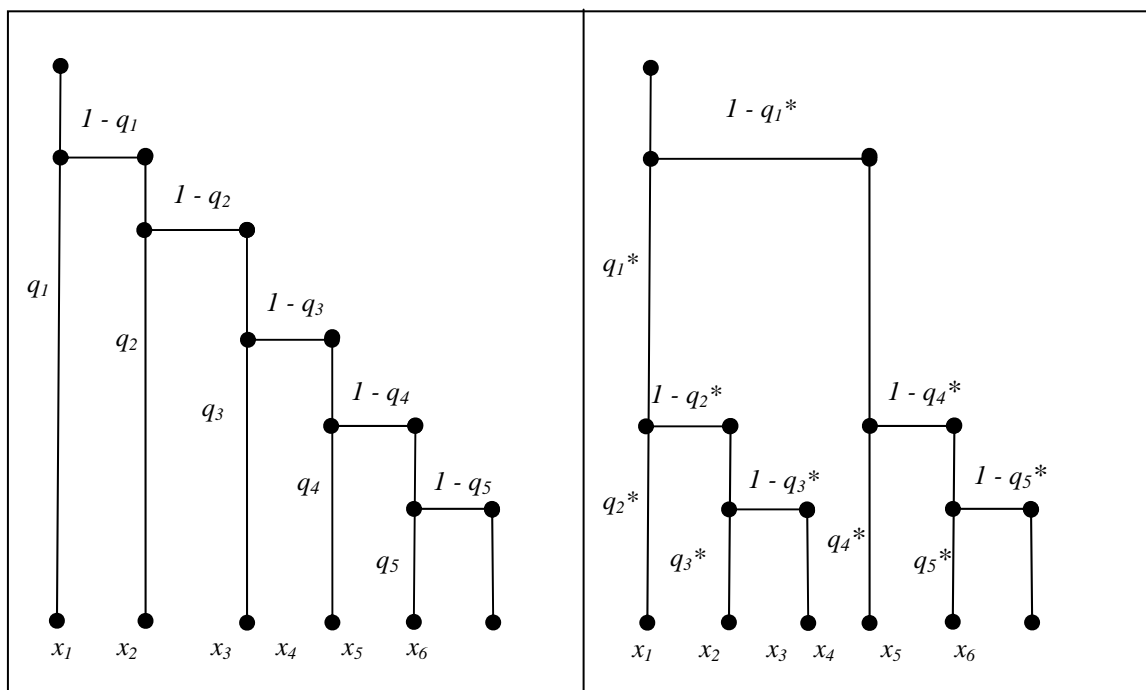
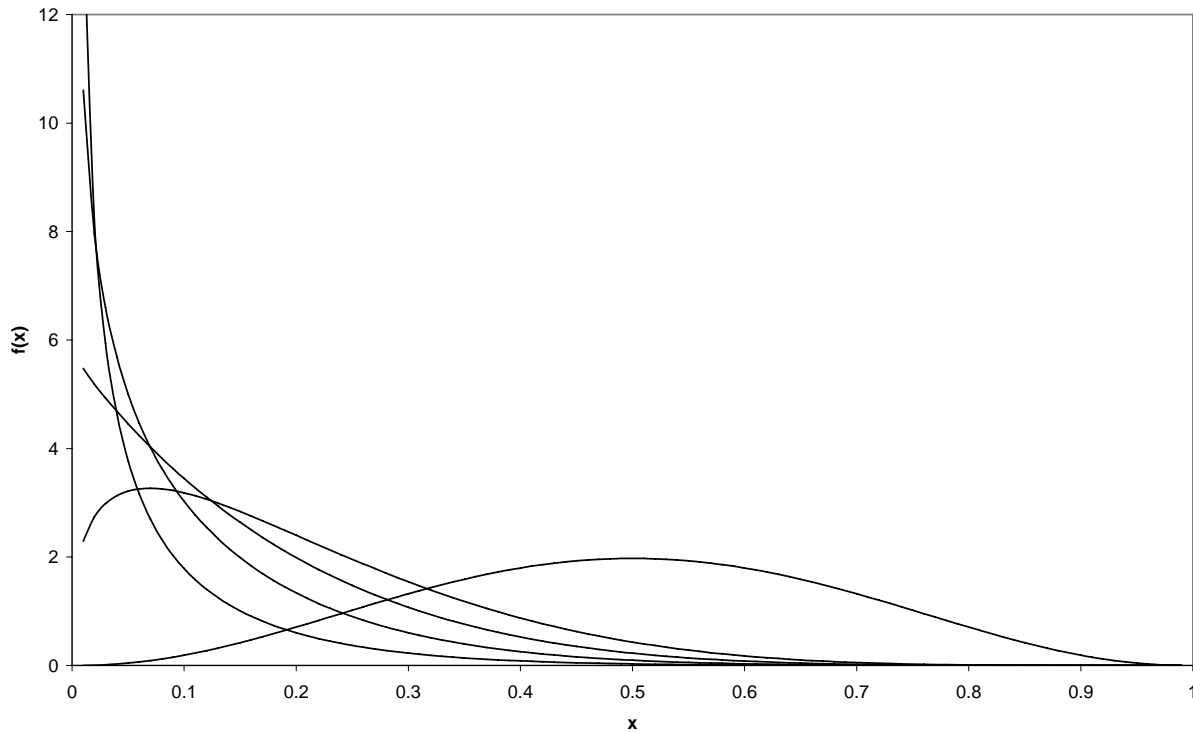


Figure 1 shows a binary bifurcation of a split fraction in the left panel and a more general decomposition in the right panel. Both types of decompositions can be accommodated in the theory and tools of this project. Figure 2 shows a set of beta marginal densities for split fraction *termini* generated through the elicitation protocol.

Figure 2

Marginal Beta Densities

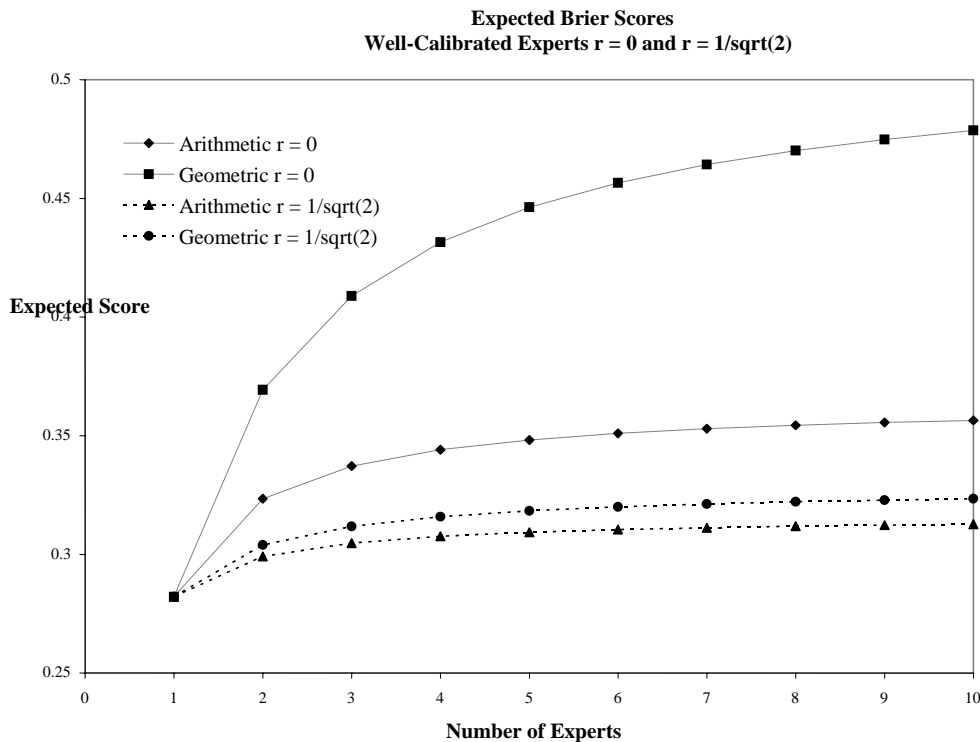


2.2. Performance of Aggregation Rules

We use infinite sequences of densities with defined properties to evaluate the expected performance of mathematical aggregation rules for elicited densities. The performance of these rules is measured through the expected variance, calibration, and expected Brier score of the aggregate. A general result for the calibration of the arithmetic average of densities from well-calibrated independent experts is given. Arithmetic and geometric aggregation rules are compared in several demonstrations using sequences of uniform, normal, and exponential densities. Sequences are developed that exhibit dependence among experts and lack of calibration. The impact of correlation, number of experts, and degree of calibration on the performance of the aggregation is demonstrated with normal densities. Performance is measured by calibration, the expected variance of the aggregate, and the expected Brier score of the aggregate.

The Figure 3 shows how the expected Brier score varies with the number of experts and the dependence among experts giving normal distributions when arithmetic or geometric aggregation is used. The graph is representative of the findings that are a result of the proposed methodology.

Figure 3



2.3. Calibration of Multivariate Densities

In this work, we push forward the understanding of probability elicitation for multivariate densities. Specifically, we have developed:

- A definition of calibration for multivariate densities (complete calibration)
- Calibration traces that graphically depict calibration
- A necessary condition for complete calibration
- Diagnostic tools for assessing calibration

Calibration is fundamental concept in probability assessment. It refers to the authenticity of the information provided by expert in the sense that the probabilities are empirically verifiable. We refer to complete calibration as the multivariate analog to the univariate concept of being well-calibrated. For a sequence of k -variate densities $f_i(\mathbf{y})$ and realizations or true values \mathbf{x}_i , we have

Definition: *The sequence $[f_i(\mathbf{y}), \mathbf{x}_i], i = 1, \dots$ is completely calibrated if the corresponding univariate marginal and conditional distribution functions are uniformly calibrated.*

From this definition, we derive calibration traces for multivariate densities and provide a necessary condition for complete calibration. Work is underway to identify the companion sufficient conditions. These tools are useful in the assessment of the quality of expert judgments and for understanding the benefits and drawbacks of various aggregation methods and how these methods impact calibration of the aggregate.

2.4. Probabilistic Modeling

In this project, we consider the situation where there are two participants -- a terrorist group and a government being targeted by the terrorist group. Both participants may allocate resources to enhance the likelihood of obtaining certain goals. For the terrorists, the goal is to inflict damage on the government while government's goal is to prevent that damage. The benefits/costs are asymmetric in that the damage to the government does not necessarily equal the benefit to the terrorists. There exists a finite set of activities in which the terrorists may engage and the government may defend itself. Each participant may invest resources that will impact the probability of terrorist act being successful. The probability of a success attack is an increasing function of the terrorist investment and a decreasing function of the investment by the government.

In the simplest version of the model, it is assumed that investment impacts only the probability of success and not the degree of damage. Thus, for the government, we consider only deterrence and not mediation of the consequences. It may also be assumed the investment by the government is public information so that the terrorists may optimize their investment using knowledge of the government investment or the government may chose to disguise its expenditures thereby misleading terrorists.

Below is the current spreadsheet interface for the implementation of this model in Excel set up to evaluate the benefit to hiding/obfuscating government expenditures to deter terrorist attacks. In this version of the model, government expenditures for various types of targets are disguised by giving out misleading information. The solution is found through Kuhn-Tucker optimization conditions.

Figure 4
Sample Spreadsheet Interface

Target	Terrorist Investment (millions \$)	Government Investment Deterrence (Billions \$)	Terrorist Benefit	Public Loss	Undisclosed expenditure % (negative is overstating)	Probability of Successful Attack
Public Buildings	0.62	3.60	10	30	0.0%	5.2%
Water supply	3.30	6.67	11	50	50.0%	3.2%
Military bases	0.96	4.59	20	40	0.0%	4.0%
Rail system	0.75	2.87	10	15	-25.0%	10.4%
Ports	0.29	4.22	15	100	0.0%	1.6%
Health system	0.55	4.62	20	70	0.0%	2.3%
National Leaders	0.97	9.25	100	1000	25.0%	0.2%
Air travel	1.94	5.91	50	50	0.0%	3.2%
Internet	0.51	4.62	20	75	0.0%	2.1%
Telecommunication	0.12	3.66	10	150	0.0%	1.0%

Allocated	10.00	50.00	5.68	15.76	Terrorist Multiplier	0.81
Budget	10	50	Total Cost	65.76	Government Multiplier	1.1022

3. Applied Relevance

3.1. Bioterrorism

Risk modeling invariably requires probabilistic quantification. Expert judgment is often relied upon when data are incomplete or conflicting so that an unambiguous quantification of the model is not forthcoming. The work on quantification of split fractions explicitly addresses a situation encountered in the NBACC modeling of bioterror threats. The findings span the range from theoretical to practical in that the theory suggests specific elicitation strategies that are immediately applicable to probability assessment of split fraction models.

3.2. Aggregation of Expert Judgments

In risk and decision analyses, including analyses of terrorist threats, multiple experts are often engaged in the quantification process. When these experts produce judgments about the same issues and quantities, it may be desirable to aggregate these judgments into a single view. Aggregation of judgments may be done because it is believed the composite results will be more accurate, it may be more convenient to deal with a single probability or density than to deal individually with the judgments of multiple experts, or because presenting a single aggregated judgment avoids the prospect of having self-conflicting results. To date, most of our understanding about aggregation has been through the analysis of properties. For example, one such property is that if the order of aggregation and Bayesian revision is reversed, the resulting probability or density will be the same. Only certain rules possess this property. There is a fairly extensive list of such properties and it is known that no rule can satisfy all these properties.

We have taken a fresh approach to the evaluation of aggregation rules. We call this approach the “performance” approach as opposed the “properties” approach. This method allows the analytic evaluation of aggregation rules in terms of the performance measures such as the calibration, spread, and expected proper score of a aggregated density. In some important and interesting cases, closed-form results have been obtained. This approach greatly enhances our understanding of the performance of various aggregation methods under specified circumstances such as the degree of calibration of the experts and their interdependence.

This work is currently being extended to the multivariate domain.

3.3. Modeling Antiterrorist Budget Allocation Decisions as a Two Person Game

The practical implication of this work is to shed light on the value of concealing or disguising information about government budget allocation designed to interdict terrorist attacks. The model provides optimal allocations for a terrorist group given government allocations. By altering the information received by the terrorist, one can determine the suboptimal response from misallocation and the reduction in threat to the country. Strategies emerge in such a model such as overstating the allocation made to a vulnerable but valuable targets in order to deter terrorists, not defending a target so that it becomes the focus of terrorist attacks (sacrifice), and pretending not to defend a target while secretly bolstering defenses in the hope of drawing an attack that will fail.

4. Collaborative Projects

The work on split fractions was developed as part of support program for work with NBACC.

5. Research Products

5.1. Publications and Reports

	Ref	Not Ref.
1. Hora, S. “An Analytic Method for Evaluating the Performance of Aggregation Rules for Probability Densities,” <i>Management Science</i> (submitted for publication)	x	
2. Hora, S., “Expert Judgment in Risk Analysis,” <i>Encyclopedia of Quantitative Risk Assessment</i> , Everitt, B., Melnick, E. (eds.) John Wiley & Sons, 2008		x
3. Hora, S., Sallaberry, C., Helton, J., “Extension of Latin Hypercube Samples with Correlated Variables,” <i>Reliability Engineering & System Safety</i> , 93 (7), Vol 93:7, 1047-1059, 2008	x	
4. Hora, S., “Expert Judgment, in <i>Advances in Decision Analysis: From Foundations to Applications</i> ,” Edwards, W., von Winterfeldt, D., Miles, R., (eds.) Cambridge University Press, 2007		x

5.2. Presentations

The Elicitation of Uncertainties over Split Fractions: An Application to Terrorism presented at the *Western Regional Science Association Annual Meeting*, Waikoloa, Hawaii, February 20, 2008.

Mathematical Aggregation of Multivariate Expert Densities presented at the *XIII International Conference on the Foundations and Applications of Utility, Risk and Decision Theory FUR XIII*, Barcelona July 2, 2008

5.3. Tools

The findings from the work on split fractions have been integrated into an elicitation tool by Dr. Tom Eppel, UCI. The work on terrorism as a two person game has produced a spreadsheet model with a variable number of targets and the ability for the government to disguise its allocations and mislead the terrorist group.