Uncertainty Analysis and Dependence Modeling

Wags and Bogsats[a]

". . .whether true or not [it] is at least probable; and he who tells nothing exceeding the bounds of probability has a right to demand that they should believe him who cannot contradict him". Samuel Johnson, author of the first English dictionary, wrote this in 1735. He was referring to the Jesuit priest Jeronimo Lobos' account of the unicorns he saw during his visit to Abyssinia in the seventeenth century [1, p. 200].

Johnson could have been the apologist for much of what passed as decision support in the period after World War II, when think tanks, forecasters, and expert judgment burst upon the scientific stage. Most salient in this genre is the book *The Year 2000* [2], in which the authors published 25 "even money bets" predicting features of the year 2000, including interplanetary engineering and conversion of humans to fluid breathers. Essentially, these are statements without pedigree or warrant, whose credibility rests on shifting the burden of proof. Their cavalier attitude toward uncertainty in quantitative decision support is representative of the period. Readers interested in knowing how many of these even money bets the authors have won, and in other examples from this period are referred to [3, Chapter 1].

Quantitative models pervade all aspects of decision making, for example, failure probabilities of unlaunched rockets, risks of nuclear reactors, effects of pollutants on health and the environment, or consequences of economic policies. Such quantitative models generally require values for parameters that cannot be measured or assessed with certainty. Engineers and scientists sometimes cover their modesty with churlish acronyms designating the source of ungrounded assessments. Wild-ass guess "(Wags)" and bunch of guys sitting around a table "(bogsats)" are two examples found in published documentation.

Decision makers, especially those in the public arena, increasingly recognize that input to quantitative models is uncertain, and demand that this uncertainty be quantified and propagated through the models (*see* **Decision Modeling**).

Initially, the modelers were the ones who provided assessments of uncertainty and did the propagating. Not surprisingly, this activity was considered secondary to the main activity of computing "nominal values" or "best estimates" to be used for forecasting and planning, and received cursory attention.

Figure 1 shows the result of such in-house uncertainty analysis performed by the National Radiological Protection Board (NRPB) and The Kernforschungszentrum Karlsruhe (KFK) in the late 1980s [4, 5]. The models in question predict the dispersion in the atmosphere of radioactive material following an accident with a nuclear reactor. The figure shows predicted lateral dispersion under stable conditions, and also shows wider and narrower plumes, which the modelers are 90% certain will enclose an actual plume under the stated conditions.

It soon became evident that if things were uncertain, then experts might disagree, and using one expert-modeler's estimates of uncertainty might not be sufficient. Structured expert judgment (*see* **Expert Judgment**) has since become an accepted method for quantifying models with uncertain input. "Structured" means that the experts are identifiable, the assessments are traceable, and the computations are transparent. To appreciate the difference between structured and unstructured expert judgment, Figure 2 shows the results of a structured expert judgment quantification of the same uncertainty pictured in Figure 1 [6]. Evidently, the picture of uncertainty emerging from these two figures are quite different.

One of the reasons for the difference between these figures is the following: the lateral spread of a plume as a function of down wind distance x is modeled, per stability class, as

$$\sigma(x) = Ax^B \qquad (1)$$

Both the constants A and B are uncertain, as attested by spreads in published values of these coefficients. However, these uncertainties can *not* be independent. Obviously, if A takes a large value, then B will tend to take smaller values. Recognizing the implausibility of assigning A and B as *independent* uncertainty distributions, and the difficulty of assessing a joint distribution on A and B, the modelers elected to consider B as a constant; that is, as known with certainty.[b]

The differences between these two figures reflect a change in perception regarding the goal of quantitative modeling. With the first picture the main effort
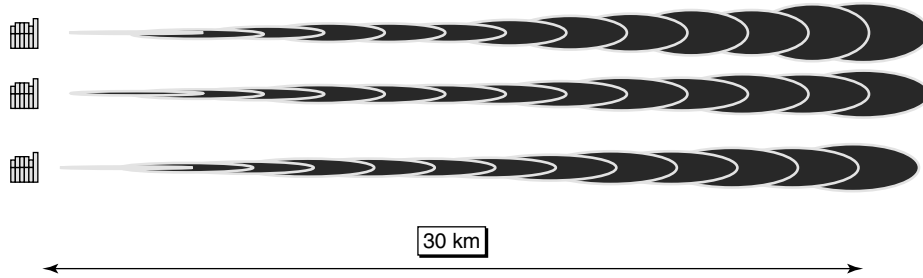
Figure 1   Plume widths (stability D) of 5, 50, and 95% computed by NRPB and KFK
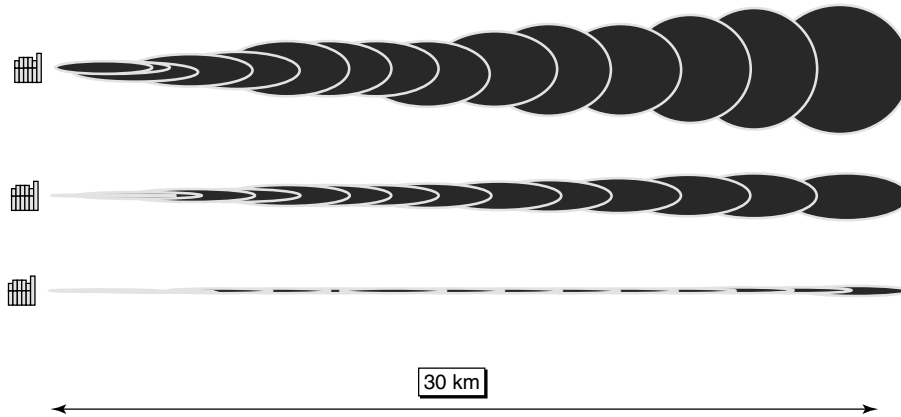

Figure 2   Plume widths (stability D) of 5, 50, and 95% computed by the EU-USNRC Uncertainty Analysis of accident consequence codes

has gone into constructing a quantitative deterministic model, to which uncertainty quantification and propagation are added on; however, in the second picture, the model is essentially about capturing uncertainty. Quantitative models are useful insofar as they help us resolve and reduce uncertainty. Three major differences in the practice of quantitative decision support follow from this shift in perception.

- First of all, the representation of uncertainty *via* expert judgment, or some other method is seen as a scientific activity subject to methodological rules every bit as rigorous as those governing the use of measurement or experimental data.
- Second, it is recognized that an essential part of uncertainty analysis is the analysis of dependence. Indeed, if all uncertainties are independent, then their propagation is mathematically trivial (though perhaps computationally challenging). Sampling and propagating independent

uncertainties can easily be trusted to the modelers themselves. However, when uncertainties are dependent, things become much more subtle, and we enter a domain for which modelers' training has not prepared them.
- Finally the domains of communication with the problem owner, model evaluation, etc. undergo significant transformations, once we recognize that the main purpose of models is to capture uncertainty.

Uncertainty Analysis and Decision Support: a Recent Example

A recent example serves to illustrate many of the issues that arise in quantifying uncertainty for decision support. The example concerns transport of Campylobacter infection in chicken processing lines.

The intention here is not to understand Campylobacter infection, but to introduce topics later in this entry. For details on Campylobacter, see [7–9].

Campylobacter contamination of chicken meat may be responsible for up to 40% of Campylobacter associated gastroenteritis and for a similar proportion of deaths. A recent effort to rank various control options for Campylobacter contamination has led to the development of a mathematical model of a processing line for chicken meat (these chickens are termed *broilers*).

A typical broiler processing line involves a number of phases as shown in Figure 3. Each phase is characterized by transfers of Campylobacter colony forming units from the chicken surface to the environment, from the environment back to the surface, from the feces to the surface (until evisceration), and the destruction of the colonies. The general model, applicable with variations in each processing phase, is shown in Figure 4.
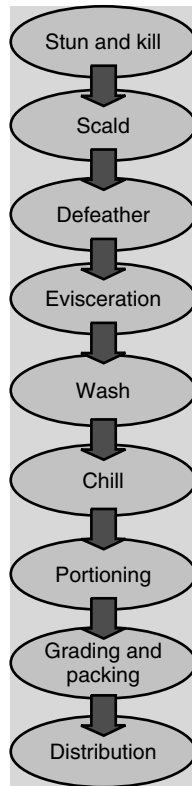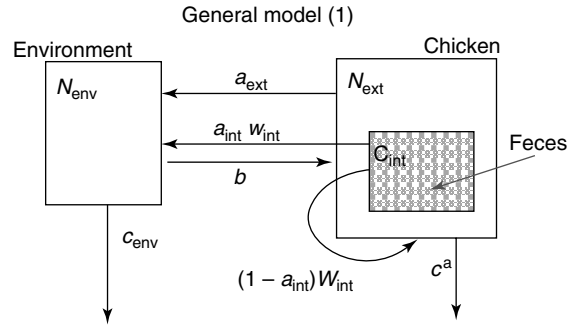

Figure 4   Transfer coefficients in a typical phase of a broiler chicken processing line

Given the number of campylobacter on and in the chickens at the inception of processing, and given the number initially in the environment, one can run the model with values for the transfer coefficients and compute the number of campylobacter colonies on the skin of a broiler and in the environment at the end of each phase. Ideally, we would like to have field measurements or experiments to determine values for the coefficients in Figure 4. Unfortunately, these are not feasible. Failing that, we must quantify the *uncertainty* in the transfer coefficients, and propagate this uncertainty through the model to obtain uncertainty distributions on the model output.

This model has been quantified in an expert judgment study involving 12 experts [9]. Methods for applying expert judgments are reviewed in [3, 10]. We may note here that expert uncertainty assessments are regarded as statistical hypotheses which may be tested against data and combined with a view to optimize performance of the resulting "decision maker".

The experts have detailed knowledge of processing lines; but, owing to the scarcity of measurements, they have no direct knowledge of the transfer mechanisms defined by the model. Indeed, use of environmental transport models is rather new in this area, and unfamiliar. Uncertainty about the transfer mechanisms can be large; and, as in the dispersion example discussed above, it is unlikely that these uncertainties could be independent. Combining possible values for transfer and removal mechanism independently would not generally yield a plausible picture. Hence, uncertainty in one transfer mechanism cannot be addressed independently of the rest of the model.

Figure 3   Broiler chicken processing line

Our quantification problem has the following features:

- There are no experiments or measurements for determining values.
- There is relevant expert knowledge, but it is not directly applicable.
- The uncertainties may be large and may not be presumed to be independent, and hence dependence must be quantified.

These obstacles will be readily recognized by anyone engaged in mathematical modeling for decision support beyond the perimeter of direct experimentation and measurement. As the need for quantitative decision support rapidly outstrips the resources of experimentation, these obstacles must be confronted and overcome. The alternative is regression to wags and bogsats.

Although experts cannot provide useful quantification for the transfer coefficients, they are able to quantify their uncertainty regarding the number of campylobacter colonies on a broiler in the situation described below taken from the elicitation protocol:

At the beginning of a new slaughtering day a thinned-flock is slaughtered in a 'typical large broiler chicken slaughterhouse'. ... We suppose every chicken to be externally infected with 10^5 campylobacters per carcass and internally with 10^8 campylobacters per gram of caecal content at the beginning of each slaughtering stage.... Question A1: All chickens of the particular flock are passing successively each slaughtering stage. How many campylobacters (per carcass) will be found after each of the mentioned stages of the slaughtering process, each time on the first chicken of the flock?

Experts respond to questions of this form, for different infection levels, by stating the 5, 50, and 95% quantiles, or percentiles of their uncertainty distributions. If distributions on the transfer coefficients in Figure 4 are given, then distributions, per processing phase, for the number of campylobacter per carcass (the quantity assessed by the experts) can be computed by Monte Carlo simulation: we sample a vector of values for the transfer coefficients, compute a vector of campylobacter per carcass, and repeat this until suitable distributions are constructed. We would like the distributions over the assessed quantities computed in this way to agree with the quantiles given by the combined expert assessments. Of course, we could guess an initial distribution over

the transfer coefficients, perform this Monte Carlo computation, and see if the resulting distributions over the assessed quantities happen to agree with the experts' assessments. In general, they will not, and this trial-and-error method is quite unlikely to produce agreement. Instead, we start with a diffuse distribution over the transfer coefficients, and adapt this distribution to fit the requirements in a procedure called *probabilistic inversion*.

More precisely, let X and Y be n- and m-dimensional random vectors, respectively, and let G be a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. We call $x \in \mathbb{R}^n$ an inverse of $y \in \mathbb{R}^m$ under G if $G(x) = y$. Similarly, we call X a probabilistic inverse of Y under G if $G(X) \sim Y$, where $\sim$ means 'has the same distribution as'. If $\{Y | Y \in C\}$ is the set of random vectors satisfying constraints C, then we say that X is an element of the probabilistic inverse of $\{Y | Y \in C\}$ under G if $G(X) \in C$. Equivalently, and more conveniently, if the distribution of Y is partially specified, then we say that X is a probabilistic inverse of Y under G if $G(X)$ satisfies the partial specification of Y. In the current context, the transfer coefficients in Figure 4 play the role of X, and the assessed quantities play the role of Y.

In our campylobacter example, the probabilistic inversion problem may now be expressed as follows: find a joint distribution over the transfer coefficients, such that the quantiles of the assessed quantities agree with the experts' quantiles. If more than one such joint distribution exists, pick the least informative of these. If no such joint distribution exists, pick a "best fitting" distribution, and assess its goodness of fit.

In fact, the best fit produced with the model in Figure 4 was not very good. It was not possible to find a distribution over the transfer coefficients which, when pushed through the model, yielded distributions matching those of the experts. Reviewing the experts' reasoning, it was found that the "best" expert in fact recognized two types of transfer from the chicken skin to the environment. A rapid transfer applied to campylobacter on the feathers, and a slow transfer applied to campylobacter in the pores of the skin. When the model was extended to accommodate this feature, a satisfactory fit was found. The second model, developed after the first probabilistic inversion, is shown in Figure 5.

Distributions resulting from probabilistic inversion typically have dependencies. In fact, this is one of the
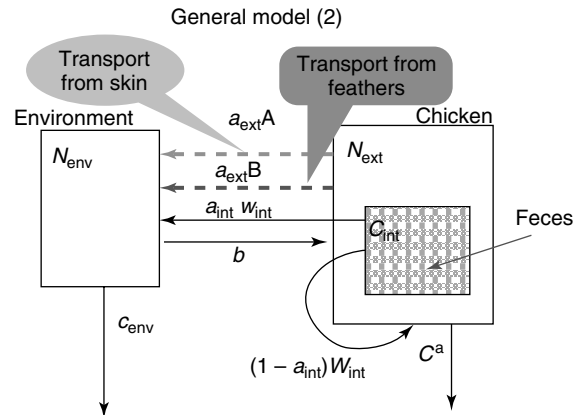
Figure 5 Processing phase model after probabilistic inversion

ways in which dependence arises in uncertainty analysis. We require tools for studying such dependencies. One simple method is simply to compute rank correlations. Rank correlation is the correlation of the quantile transforms of the random variables, where the quantile transform simply replaces the value of a random variable by its quantile. There are many reasons for preferring the rank correlation to the more familiar product moment correlation in uncertainty analysis.[c] For now, it will suffice simply to display in Table 1 the rank correlation matrix for the transfer coefficients in Figure 5, for the scalding phase.

Table 1 shows a pronounced negative correlation between the rapid transfer from the skin (a_{extA}) and evacuation from the chicken (ca), but other correlations are rather small. Correlations do not tell the whole story; they are averages after all. To obtain a detailed picture of interactions, other tools must be applied. One such tool is the cobweb plot. In a cobweb plot, variables are represented as vertical lines. Each sample realizes one value

of each variable. Connecting these values by line segments, one sample is represented as a jagged line intersecting all the vertical lines. Figure 6 shows 2000 such jagged lines and gives a picture of the joint distribution. In this case we have plotted the quantiles, or percentiles, or ranks of the variables rather than the values themselves. Contracting the name of a_{extA} to axa, the negative rank correlation between axa and ca is readily visible if the picture is viewed in color: the lines hitting low values of a_{extA} are red, and the lines hitting high values of ca are also red.

We see that the rank dependence structure is quite complex. Thus, we see that low values of the variable ce (c_{env}, the names have been shortened for this graph) are strongly associated with high values of b, but high values of ce may occur equally with high and low values of b. Correlation (rank or otherwise) is an average association over all sample values and may not reveal complex interactions that are critical for decision making. One simple illustration highlights their use in this example.

Suppose we have a choice of accelerating the removal from the environment ce or from the chicken ca; which would be more effective in reducing campylobacter transmission? To answer this, we add two output variables: $a1$ (corresponding to the elicitation question given above) is the amount on the *first* chicken of the flock as it leaves the processing phase, and $a2$ is the amount on the *last* chicken of the flock as it leaves the processing phase. In Figure 7, we have conditionalized the joint distribution by

Table 1 Rank correlation matrix of transfer coefficients, scalding phase

Variable	a_{extA}	a_{extB}	ca	b	ce	$aint$
a_{extA}	1.00	0.17	−0.60	−0.04	0.03	0.00
a_{extB}	0.17	1.00	−0.19	−0.10	−0.06	0.00
ca	−0.60	−0.19	1.00	0.01	0.02	0.00
b	−0.04	−0.10	0.01	1.00	0.02	0.00
ce	0.03	−0.06	0.02	0.02	1.00	0.00
$aint$	0.00	0.00	0.00	0.00	0.00	0.00

Samples selected: 2000
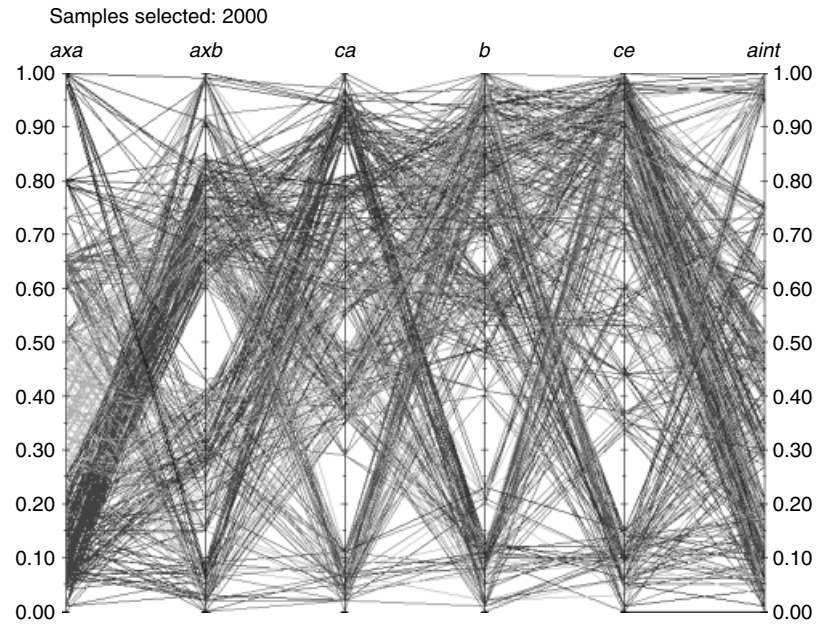
Figure 6   Cobweb plot for the transfer coefficients in the extended model
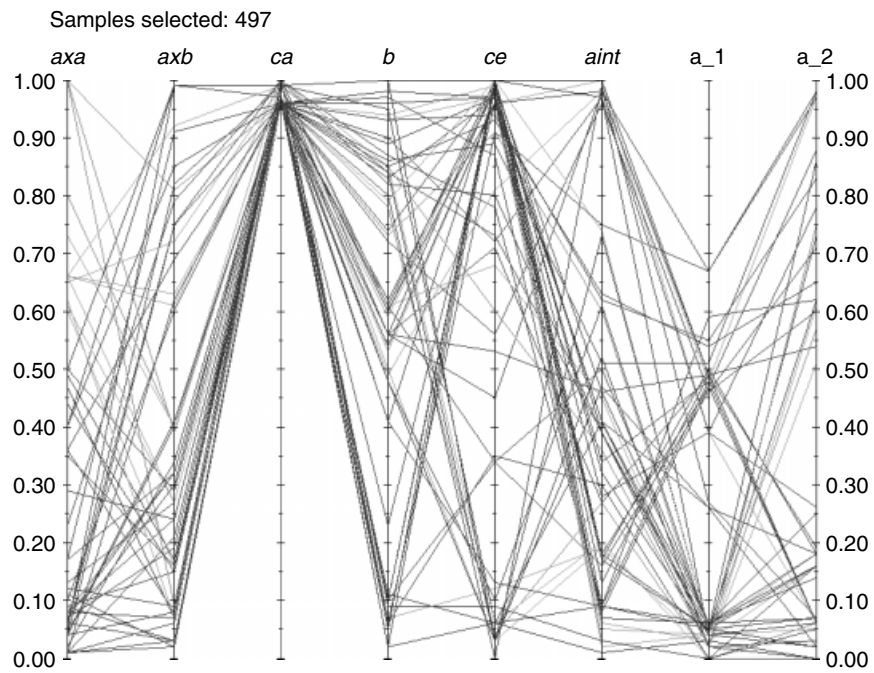
Samples selected: 497

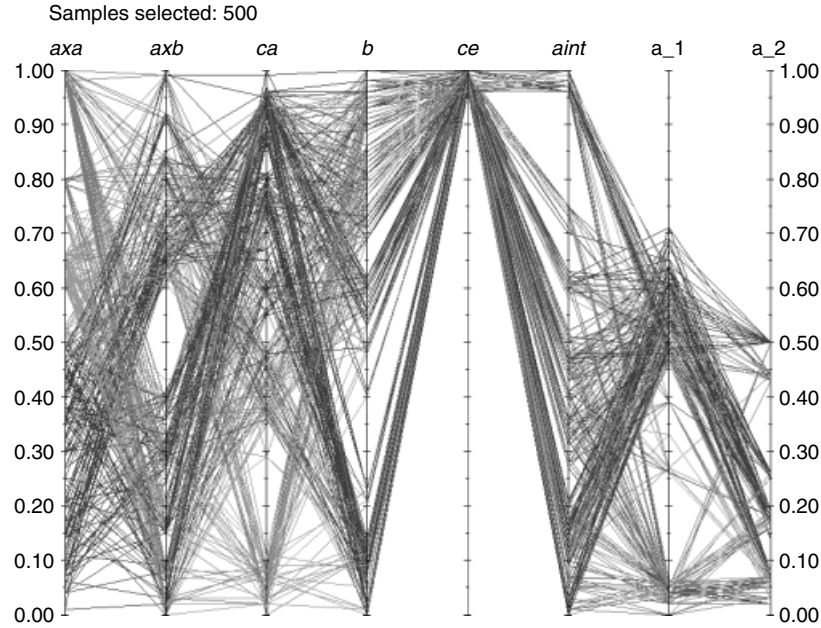**Figure 7**   Cobweb plot conditional on high *ca*

Figure 8 Cobweb plot conditional on high *ce*

selecting the upper 5% of the distribution for *ca*; in Figure 8, we do the same for *ce*.

Let us assume that the intervention simply conditionalizes our uncertainty, without changing other causal relations in the processing line (otherwise, we should have to rebuild our model). On this assumption, it is readily apparent that *ce* is more effective than that on *ca*, especially for the last chicken.

This example illustrates a feature that pervades quantitative decision support, namely, that input parameters of the mathematical models cannot be known with certainty. In such situations, mathematical models should be used to capture and propagate uncertainty. They should not be used to help a bunch of guys sitting around a table make statements that should be believed if they cannot be contradicted. In particular, it shows that

- Expert knowledge can be brought to bear in situations where direct experiment or measurement is not possible, namely, by quantifying expert uncertainty on variables that the models should predict.
- Utilizing techniques like probabilistic inversion in such situations, models become vehicles for capturing and propagating uncertainty.

- Configured in this way, expert input can play an effective role in evaluating and improving models.
- Models quantified with uncertainty, rather than wags and bogsats, can provide meaningful decision support.

Uncertainty Analysis – the State of the Art

The following remarks focus on techniques for uncertainty analysis that are generally applicable. This means, uncertainty distributions may *not* be assumed to conform to any parametric form and techniques for specifying, sampling, and analyzing high dimensional distributions should therefore be nonparametric.

Some techniques, in particular, those associated with bivariate dependence modeling, are becoming familiar to a wide range of users. Copulae, or multivariate distributions with uniform margins are used to capture rank correlation structure which is independent of marginal distributions (*see* **Copulas and Other Measures of Dependency**). While two-dimensional copulae are becoming familiar, multidimensional copulae remain extremely limited in their ability to represent dependence. Good books are available for bivariate dependence, such as [11–14].

High dimensional dependence models, sampling methods, postprocessing analysis, and probabilistic inversion are "breaking news" to nonspecialists, both mathematicians and modelers.

With regard to dependence in higher dimensions, much is not known. For example, we do not know whether an arbitrary correlation matrix is also a rank correlation matrix.[d] We do know that characterizing dependence in higher dimensions *via* product moment correlation matrices is *not* the way to go. Product moment correlations impose unwelcome constraints on the one-dimensional distributions. Further, correlation matrices must be positive definite, and must be completely specified. In practice, data errors, rounding errors, or simply vacant cells lead to intractable problems with regard to positive definiteness. We must design other, friendlier ways to let the world tell us, and to let us tell computers which high dimensional distribution to calculate. We take the position that graphical models are the weapon of choice. These may be Markov trees, vines, independence graphs or Bayesian belief nets. For constructing sampling routines capable of realizing richly complex dependence structures, we advocate regular vines. They also allow us to move beyond discrete Bayesian belief nets without defaulting to the joint normal distribution. Much of this material is new and only very recently available in the literature [15–20].

Problems in measuring, inferring, and modeling high dimensional dependencies are mirrored at the end of the analysis by problems in communicating this information to problem owners and decision makers. This is sometimes called *sensitivity analysis* [21].

Whereas, communicating uncertainty has received some attention, much less attention has gone into *utilizing* uncertainty. It is safe to say that our ability to quantify and propagate uncertainty far outstrips our ability to *use* this quantification to our advantage in structured decision problems.

End notes

[a.] This chapter is based on the first Chapter of [1] to which the reader may refer for definitions and details.
[b.] This is certainly not the only reason for the differences between Figures 1 and 2. There was also ambivalence with regard to what the uncertainty should capture. Should it capture the plume uncertainty in a single accidental release, or the

uncertainty in the average plume spread in a large number of accidents? Risk analysts clearly required the former, but meteorologists are more inclined to think in terms of the latter.
[c.] Among them the rank correlation always exists, is independent of the marginal distributions, and is invariant under monotonic increasing transformations of the original variables.
[d.] We have recently received a manuscript from H. Joe that purports to answer this question in the negative for dimension greater than four.

References

[1] Shepard, O. (1930). *The Lore of the Unicorn*, Avenel Books, New York.

[2] Kahn, H. & Wiener, A.J. (1967). *The Year 2000: A Framework for Speculation*, Macmillan, New York.

[3] Cooke, R.M. (1991). *Experts in Uncertainty*, Oxford University Press, New York.

[4] Fischer, F., Ehrhardt, J. & Hasemann, I. (1990). Uncertainty and sensitivity analyses of the complete program system ufomod and selected submodels, Technical Report 4627, Kernforschungszentrum Karlsruhe.

[5] Crick, J.J., Hofer, E., Johnes, J.A. & Haywood, S.M. (1988). Uncertainty analysis of the foodchain and atmospheric dispersion modules of MARC, Technical report NRBP-R184, National Radiological Protection Board, Chilton, Didcot, Oxon.

[6] Cooke, R.M. (1997). Uncertainty modeling: examples and issues, *Safety Science* **26**(1/2) 49–60.

[7] Nauta, M.J., van der Fels-Klerx, H.J. & Havelaar, A.H. (2004). A poultry processing model for quantitative microbiological risk assessment, *Risk Analysis* (submitted).

[8] Cooke, R.M., Nauta, M., Havelaar, A.H. & van der Fels, I. Probabilistic inversion for chicken processing lines, *Reliability Engineering and System Safety* (in press).

[9] van der Fels-Klerx, H.J., Cooke, R.M., Nauta, M.J., Goossens, L.H.J. & Havelaar, A.H. (2005). A structured expert judgment study for a model of campylobacter contamination during broiler chicken processing, *Risk Analysis* **25**, 109–124.

[10] Cooke, R.M. & Goossens, L.H.J. (2000). Procedures guide for structured expert judgement, Technical Report EUR 18820 EN, European Commission, Directorate-General for Research, Nuclear Science and Technology, Brussels.

[11] Doruet Mari, D. & Kotz, S. (2001). *Correlation and Dependence*, Imperial College Press, London.

[12] Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.

[13] Nelsen, R.B. (1999). *An Introduction to Copulas*, Springer, New York.

[14] Dall'Aglio, G., Kotz, S. & Salinetti, G. (1991). *Probability Distributions with Given Marginals; Beyond the Copulas*, Kulwer Academic Publishers.

[15] Bedford, T.J. & Cooke, R.M. (2001). *Probabilistic Risk Analysis; Foundations and Methods*, Cambridge University Press, Cambridge.

[16] Bedford, T.J. & Cooke, R.M. (2002). Vines – a new graphical model for dependent random variables, *Annals of Statistics* **30**(4), 1031–1068.

[17] Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*, John Wiley & Sons, Chichester.

[18] Cowell, R.G., Dawid, A.P., Lauritzen, S.L. & Spiegelhalter, D.J. (1999). Probabilistic networks and expert systems, in *Statistics for Engineering and Information Sciences*, Springer-Verlag, New York.

[19] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufman Publishers, San Mateo.

[20] Kurowicka, D. & Cooke, R.M. (2006). *Uncertainty Analysis with High Dimensional Dependence Modeling*, John Wiley & Sons, New York.

[21] Saltelli, A., Chan, K. & Schott, E. (2000). *Sensitivity Analysis*, John Wiley & Sons, New York.

Related Articles

Bayesian Statistics in Quantitative Risk Assessment

Decision Analysis

Group Decision

Societal Decision Making

ROGER M. COOKE AND DOROTA KUROWICKA