Multivariate Elicitation

Previous versions of SHELF have concentrated on eliciting expert knowledge about one uncertain quantity at a time, and much of the material in the SHELF package is aimed at this task. This includes all of the elicitation methods discussed in the “SHELF Methods” document, and the “SHELF 2 (Distribution)” template. Eliciting probability distributions for two or more uncertain quantities together introduces additional complications unless the quantities can be considered independent. In this document, we first define independence and discuss the implications for elicitation of dependence. We then consider the possibility of removing dependence through elaboration, and finally present the two multivariate elicitation methods that are implemented in the new templates “SHELF 3 (Multivariate) C” and “SHELF 3 (Multivariate) D”.

Independence

A set of quantities are independent if learning additional information about one of them would not affect an expert’s judgements about the others. Notice that independence is itself a subjective judgement. For instance, two quantities X and Y may be independent for one expert but not for another.

If *all* the experts judge that two quantities X and Y are independent then it is enough to elicit a probability distribution for each quantity separately. We do not need to consider what is known about Y when eliciting a distribution for X, because knowledge about Y has no bearing on the experts’ judgements about X, and conversely we do not need to consider knowledge about X when eliciting their distribution for Y.

If X and Y are *not* judged to be independent, then we can still elicit distributions for X and Y separately, but these *marginal* distributions are now not a complete description of the experts’ knowledge about X and Y. We also need to elicit *how* knowledge of one quantity would affect judgements about the other.

To see why this is necessary, suppose that X and Y are two uncertain inputs to a risk model, and suppose that larger values of either X or Y lead to increased risk. If X and Y are independent, then a large value of X does not make a large value of Y more or less likely, but this is not the case when we have dependence. It may be that the experts judge that if X is large then Y is likely also to be large. This is called positive dependence, and would imply an expectation of greater risk than would exist in the case of independence. Conversely, negative dependence, i.e. larger values of one quantity making large values of the other *less* likely, would imply less risk overall than the case of independence.

In general, when some quantities are dependent, what is needed is their *joint* distribution. Whereas the marginal distribution for a single quantity X specifies probabilities like $P(X > 2)$ or $P(20 < X < 24)$, the joint distribution for two quantities X and Y will specify probabilities like $P(X > 2$ *and* $Y > 3)$, $P(20 < X < 24$ *and* $-1 < Y < 0)$ or $P(X < Y)$. With independence, joint probabilities are products of marginal probabilities, for instance $P(X > 2$ *and* $Y > 3) = P(X > 2) \times P(Y > 3)$, which is why it is sufficient to elicit only marginal distributions in this case. Joint distributions are much more complex than marginal distributions, and in principle are therefore much more complex to elicit.

Elaboration

The idea of elaboration is introduced in the document "Definitions", where its role is to facilitate more accurate elicitation of a distribution for a single QoI, by expressing it in terms of two or more quantities that may individually be easier for experts to make judgements about. We now consider a broader definition of elaboration for the purposes of multivariate elicitation.

Let $\mathbf{X} = (X_1, X_2, \ldots)$ be a set of quantities of interest, which we suppose are dependent. An elaboration of $\mathbf{X}$ expresses it in terms of another set of quantities $\mathbf{Y} = (Y_1, Y_2, \ldots)$, so that formally each X_i is a function of $\mathbf{Y}$. Given a joint distribution for $\mathbf{Y}$, we can infer the joint distribution of $\mathbf{X}$. The elaboration will be successful in facilitating the task of eliciting a distribution for $\mathbf{X}$ if the experts now judge the components of $\mathbf{Y}$, i.e. $Y_1, Y_2, \ldots$, to be independent.

One approach to multivariate elicitation of dependent quantities, therefore, is to elaborate them in terms of new quantities that the experts judge to be independent. The following examples will illustrate the technique.

Example 1: Two treatment effects

Suppose that a clinical trial is to be conducted in China to compare the effect of a new drug D with a comparator drug C for treating patients with a certain disease. The quantities of interest are the mean effects of the two drugs, X_D and X_C. Both quantities are uncertain for Chinese patients, although C is the standard treatment in Europe and North America, where there is substantial evidence regarding its mean effect. Experts do not regard X_D and X_C as independent; for instance, if the effect X_C in Chinese patients is higher than in European patients, this would increase the experts' expectation of the new drug's effect X_D.

If, however, we define $Y_D = X_D/X_C$ to be the *relative* effect of drug D, relative to the comparator C, then the experts may regard this as independent of X_C. Formally, we are expressing $\mathbf{X} = (X_D, X_C)$ in terms of $\mathbf{Y} = (Y_D, Y_C)$ by the equations $Y_D = X_D/X_C$ and $Y_C = X_C$. Independence between Y_D and Y_C means that we can separately elicit distributions for these two quantities, and this will induce the required distribution of X_D and X_C. In particular, larger values of $X_C = Y_C$ will lead to expectation of larger values of X_D through the inverse relationship $X_D = Y_D \times Y_C$.

Example 2: Tensile strengths

Metal rods are available from five different manufacturers. Their manufacturing techniques are different, and the rods they produce are likely to have somewhat different properties. The quantities of interest are the mean tensile strengths X_1, X_2, X_3, X_4 and X_5 of the rods from each manufacturer. The experts judge these to be dependent in a similar way to the previous example – if we learnt the tensile strength X_1 of rods from the first manufacturer, this would influence the experts' judgements about X_2, X_3, X_4 and X_5. For instance, if X_1 is lower than expected, this would decrease the experts' expectations for the other quantities.

We could apply a similar elaboration approach to Example 1 if we have good knowledge of the mean tensile strength of rods from one manufacturer, perhaps because this has been the previous supplier. Then we could define this as the comparator, and consider the tensile strengths for the other manufacturers *relative* to this.

However, suppose we cannot single out one manufacturer as a natural comparator in this way. In this case we may elicit judgements about each of X_1, X_2, X_3, X_4 and X_5 relative to an artificial or conceptual comparator X_0. For instance, we might define X_0 to be the mean tensile strength of rods made by a typical or average manufacturer. We thereby elaborate $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5)$ in terms of $\mathbf{Y} = (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6)$, where $Y_1 = X_1 - X_0$, …, $Y_5 = X_5 - X_0$ and $Y_6 = X_0$. Notice that instead of five quantities of interest we now have six to elicit judgements for, but the added dimensionality is a small price to pay if the experts now judge the components of $\mathbf{Y}$ to be independent. With X_0 defined in this way, the experts will judge that Y_1 to Y_5 could take both positive and negative values, perhaps with higher probabilities of positive values if the selection of the five candidate manufacturers implies that they are expected to produce stronger rods than a typical manufacturer.

It is worth emphasising that the choice of elaboration is always a matter for the facilitator's judgement, based on discussion with the project team and/or the experts themselves. In this case, we might choose to express tensile strengths relative to X_0 by ratios rather than differences if the experts would find it easier to make judgements about ratios.

We might also use a different conceptual comparator. For instance, X_0 might be the mean tensile strength of rods from a perfect or ideal manufacturer. With the elaboration $Y_1 = X_1/X_0$, …, $Y_5 = X_5/X_0$ and $Y_6 = X_0$, the experts will give distributions for Y_1 to Y_5 such that they have zero probabilities of exceeding 1.

Example 3: Toxicity

The toxicity of a river water pollutant is described by the proportion T(c) of salmon that die when exposed to the pollutant for 48 hours in concentration c. The project team are interested in the *function* T(c) for a range of concentrations c. In principle, this is a very large number of quantities of interest, and they are clearly dependent. For instance, if $T(c_1)$ and $T(c_2)$ are two points on the function with $c_2 > c_1$, then we know that $T(c_2)$ must be greater than $T(c_1)$, so acquiring additional knowledge about any of these quantities must affect judgements about the others.

In addition to dependence, this example has the added complication of the large, effectively infinite, number of quantities of interest, i.e. all points on the toxicity function T(c). An elaboration will only be usable if the number of elements of **Y** is small enough, so that it is feasible to elicit expert judgements about all of these elements. This implies the need for a dramatic reduction in dimensionality, which can only be achieved by making assumptions.

A standard mathematical form for a toxicity function is the logistic function

$$T(c) = \exp(Y_2(c - Y_1))/[1 + \exp(Y_2(c - Y_1))] ,$$

Where Y_1 and Y_2 are the logistic *parameters*. Y_1 is the value of the concentration c at which 50% of the salmon will die, and is known as the LC50, while Y_2 controls the rate at which toxicity is increasing around the LC50. If we assume that T(c) has this logistic form, then we have an elaboration of the whole function in terms of just two new quantities, Y_1 and Y_2. The experts might judge these to be independent, and hence this becomes a usable elaboration.

The assumption of a particular mathematical form for a function is obviously strong, and the assumed form must be determined through discussion with the experts. A similar elaborative solution is needed for the variability problem discussed in the “Definitions” document. There, the QoI was the value X taken by a quantity A for a random member of some population. The proposed elaboration was in terms of the physical distribution of A values in the population, which is described by the proportion F(a) of members of the population with values of A less than or equal to a. Since this is a function, we need to assume a suitable mathematical form for the physical distribution, for instance that the values of A follow a Gaussian probability distribution with parameters Y_1 equal to the population mean value and Y_2 equal to the population variance.

Gaussian copula method

When it is not possible to find an elaboration of X such that the elements of Y are judged by all the experts to be independent, it is necessary to apply a specifically multivariate elicitation method. The template “SHELF 3 (Multivariate) C” implements a method based on a form of multivariate distribution known as the Gaussian copula. It is quite generally applicable, except that it cannot be used if the possible values of one QoI depend on the values of other QoIs. An example of that kind of constraint is when the QoIs must sum to 1; the template "SHELF 3 (Multivariate) D" implements a method for this case.

The copula method allows each QoI to have its own marginal distribution, which may be of any form. So the template begins by eliciting a distribution for each QoI using the "SHELF 2 (Distribution) template". Then in order to elicit dependence between the QoIs, one extra judgement is required for each pair of QoIs. This is the *concordance probability*, which is the probability that the true values of the two QoIs will both be on the same side of their elicited medians (i.e. both above their medians or both below). Ways to make and refine this judgement are explained in the slide set "Concordance Probability".

It is possible for the experts to specify combinations of concordance probabilities that are incompatible. For instance, if they specify a high concordance probability for X_1 and X_2 and also for X_1 and X_3, then this means that these pairs of QoIs are strongly positively dependent. High values of X_1 are associated with high values of both X_2 and X_3, and there must be at least moderately strong positive dependence between X_2 and X_3. If the experts were to give a low concordance probability for X_2 and X_3, this would be infeasible. The "SHELF" software identifies incompatible combinations of judgements. In this case the facilitator will provide feedback to the experts in order for them to make appropriate revisions of their judgements.

Although such inconsistencies may be unusual in practice, the risk increases with increasing numbers of quantities. In fact, the copula method is not really practical for more than a few quantities because of the large number of concordance probabilities that must be elicited (for instance, with just six quantities there are 15 pairs for which a concordance probability must be elicited, and this is already a large task for the experts). Since the method is also relatively untried, we recommend using it with no more than 3 or 4 quantities.

The Gaussian copula is a realistic and quite flexible method to elicit a general multivariate distribution. The concordance probabilities are not easy judgements for the experts to make, but other measures of dependence are considerably more complex. Although ideally one would elicit more judgements, to characterise dependence in more depth and detail, experts would not be able to make such judgements reliably.

Dirichlet method

The template "SHELF 3 (Multivariate) D" implements a method to elicit a multivariate distribution for a set of QoIs that are constrained to sum to 1. For example, the QoIs might be the proportions of households in a district that have 0, 1, 2, 3, 4 or more than 4 televisions. Each individual proportion could in principle take any value from 0 to 1, but the sum of all six proportions must be 1. This tends to induce negative dependence. For instance, if the proportion with 0 televisions is 0.3 (30%), then none of the other five proportions can exceed 0.7, and indeed their sum must be 0.7. The larger any one proportion is, the smaller the others must be, at least in total.

The Dirichlet method requires each QoI to have a beta marginal distribution. The template therefore begins by eliciting a distribution for each QoI, using the "SHELF 2 (Distribution)" template, but with the constraint that the fitted distribution must in each case be beta.

The template then fits a Dirichlet distribution to the elicited marginal distributions. The Dirichlet family of distributions is much more restrictive than the Gaussian copula, and experts are not asked to make any additional judgements. Indeed, the combination of elicited marginal distributions will almost certainly already be incompatible with the Dirichlet distribution. The fitting process therefore identifies a compatible set of beta marginal distributions that is close to the elicited distributions, and the facilitator will provide feedback to the experts in order for them to assess whether the revised distributions are a reasonable representation of their opinions.

Example 4: Bird abundance

In this final example, the Dirichlet method is used after a preliminary elaboration. The QoIs X_1, X_2 and X_3 are the abundances (numbers per hectare) of birds of each of three woodland species in a particular forest. The experts may be expected to judge these three quantities to be dependent, but the dependence might be positive or negative. If little is known about the total number of birds, of all species, that might be found in the forest, then large values of one species might suggest that the forest can support large bird populations, and so the numbers of other species might also be large – positive dependence. However, if the experts had good knowledge of how large a total population the forest could support, then large numbers of one species will suggest low values of the others – negative dependence.

A natural elaboration in this case is first to define X_4 to be the abundance of all other species, so that $T = X_1 + X_2 + X_3 + X_4$ is the total abundance. Then we define $Y_1 = X_1/T$, $Y_2 = X_2/T$, $Y_3 = X_3/T$, $Y_4 = X_4/T$ and $Y_5 = T$. Now Y_1 to Y_3 are the proportions of the three species and Y_4 is the proportion of all other species, and so these four quantities must sum to 1 and we can elicit a joint distribution for them using the Dirichlet method. They are of course still dependent, but the experts may judge that they are independent of the total abundance Y_5. Therefore, although the elaboration has not yielded a set of new quantities that are all independent, it has achieved the desired result of making the multivariate elicitation feasible, via a Dirichlet method for Y_1 to Y_4 and an independent elicitation of a (marginal) distribution for Y_5.