Discrete Quantities

Previous versions of SHELF have been devoted to methods for eliciting expert knowledge about one or more uncertain quantities, each of which can take any value in some range. Such quantities are called *continuous*. New templates in version 4 facilitate the elicitation of judgements about a quantity that can take a small number of distinct possible values. Such quantities are called *discrete*. An event E either happens or does not happen, so it can be thought of as a discrete quantity that can only take two possible "values".

Example 1. A company is planning to bid for a contract for a building project. The company's costing department has proposed a bid price for the work, but the managing director asks for an assessment of how likely they are to win the contract at that price. Here we have a single event W, that the company will win the contract. It can be considered as a discrete quantity taking two values which we may code as 0 and 1, where $W = 1$ corresponds to winning the contract and $W = 0$ to losing it.

Example 2. Continuing Example 1, a related question might be how many other companies will bid for the contract. There are three other companies that might submit good bids for the work, so we are interested in the discrete quantity N, which is the number of those companies that will bid. The possible values of N are 0, 1, 2 and 3.

Example 3. One factor in a risk assessment for a yogurt product concerns whether the milk arriving at the factory is contaminated with E. coli. The focus may then be on the event C that an individual batch of milk is contaminated. The event has two "values" as in Example 1.

Example 4. A surgeon will perform an urgent operation on a patient with a serious heart problem. The surgeon considers that there are three principal outcomes – death, survival but with disability, and full recovery. Again, we could code this as a discrete quantity R having three possible values: $R = 0$ denotes death, $R = 1$ is survival with disability and $R = 2$ is full recovery.

The "SHELF 3 Discrete" template

When eliciting a probability distribution for a continuous quantity, the basic procedure is to elicit a small number of judgements from the experts and then to fit a suitable distribution to those judgements. The SHELF 2 template offers a variety of methods which differ in the judgements that are elicited. This approach is required because it is not feasible to elicit the very large number of judgements required to specify the probability distribution completely for a continuous quantity.

In contrast, for a discrete quantity we can simply elicit a probability for each possible value, and there is no need for a fitting step. The SHELF 3 Discrete template is therefore somewhat simpler than SHELF 2, but with the same basic structure of a round of individual judgements, followed by discussion and then group judgements from the perspective of a Rational Impartial Observer, RIO.

In the case of an event, it is only necessary to elicit a single probability because the probability of the event not occurring must be one minus the probability of it occurring. In Example 1, for instance, we only need to elicit $P(W = 1)$, the probability of winning the contract. At the individual judgements stage, each expert provides his or her own probability for the event, and in the group judgement stage the experts agree on RIO's probability.

For a quantity with more than two possible values, it is necessary to ensure that the elicited probabilities sum to one. In this case, the Roulette method (described in full in the "SHELF Methods" advice document) is recommended for individual judgements, with each possible value of the quantity being a separate 'bin'. In the group judgement stage, the facilitator will ask the experts to make a RIO probability judgement for each possible value of X. However, if X can take more than 3 or 4 possible values, simply asking for each probability in turn invites two problems. First, the probabilities may not sum to 1, in which case the facilitator will need to ask for adjustments in order to achieve a proper probability distribution.

The second problem in asking for a sequence of probabilities is anchoring. The facilitator may instead consider a partitioning approach. The possible values are first grouped into two sets and the experts asked to provide a probability for the first set (recognising that this means that the complementary probability is assigned to the other set). Then the first set is split into two subsets and the experts divide the elicited probability for the set between those subsets. The same is then done for the other set, and so on until every possible value has been assigned a probability. This approach works particularly well when there are natural or meaningful ways to partition the possible values of X.

The partitioning can be implemented by roulette. Thus, a number of probs is allocated to the first set, representing the elicited probability for this set, with the remaining probs assigned to the second set. Then as each set is further partitioned the probs already allocated to it are shared out among the subsets, so that in the end we have a roulette-style allocation of probs to each possible value of X. This makes the partition approach clearer and more visual for the experts, although it may be necessary to use a large number of probs.

The SHELF 3 Discrete template is the basic and simplest way to address elicitation of discrete quantities, but it is not the only way. In the remaining sections of this document we will find uses for SHELF 2, SHELF 3 Dirichlet and SHELF 3 Extension in the context of discrete quantities.

Repeatable quantities

In Example 3, the event C refers to contamination of a *single* batch of milk, but it is not obvious that this is the relevant quantity of interest in the risk assessment. An alternative focus of the assessment might be the quantity F_C, defined to be the *proportion* of batches of milk that are contaminated, out of all the batches that might be received at the factory over time.

Similarly, Example 4 defines R to be the outcome of the operation for the particular patient. But we could equally consider the proportions F_0, F_1 and F_2 of patients who will die, live with disability or make a full recovery, out of all patients who might be given this operation over time.

These are examples of repeatable quantities (and specifically repeatable events in the case of Example 3). A repeatable quantity takes a value for each individual in some population. In Example 3, the population is batches of milk, while in Example 4 it is patients. The quantity of interest in an elicitation may legitimately be the value for a specific individual in the population, but more usually when we have repeatable quantities we are interested in the proportion of individuals having each of the possible values.

The letter F in the above examples represents 'frequency', a word that is technically preferable to 'proportion'.[1] There is an important difference between uncertainty about the event C or the discrete quantity R and uncertainty about the corresponding frequencies F, F_0, F_1 and F_2. Uncertainty about an individual event like C is described by a single probability, the probability that the event occurs, and uncertainty about a discrete quantity like R is described by a set of probabilities for each possible value. However, frequencies are continuous quantities, taking possible values between 0 and 1, and therefore uncertainty about a frequency is described by a probability distribution over the range 0 to 1.

In Example 3, if interest really lies in the frequency F_C, then this should be elicited using the SHELF 2 template. In Example 4, if the surgeon is really interested in the frequencies F_0, F_1, F_2, the appropriate template is SHELF 3 Dirichlet, which is for multivariate elicitation of a set of quantities that must sum to one.

[1] The reason is that when the population is at least potentially infinite the relevant mathematical definition of 'proportion' is the limit of the frequency with which a value occurs in an infinite sequence of individuals.

A common question

Frequencies are also one way to look at a very common question that arises when eliciting probabilities for events or discrete quantities.

Suppose that you wish to elicit expert judgement about a single uncertain event. For example, in a risk assessment for a nuclear power plant one possible risk is a terrorist flying an aircraft into the reactor building. During the entire operating life of the reactor, what is the chance of such an event? The result of your expert elicitation should be the elicited probability. However, a probability like this can be very difficult to quantify with any degree of confidence, and a question that often arises is whether it is necessary to express uncertainty about the probability.

This question is a very common source of confusion for facilitators, experts and clients alike.

Tossing coins and drawing-pins

One reason why people often wish to express uncertainty about an elicited probability is the fact that some probability judgements can obviously be made more confidently than others. We will begin with two very simple examples to illustrate the ideas.[2]

For simplicity, again, we will imagine probability judgements being made by a single expert, perhaps in the individual judgements round of a SHELF elicitation, but the same considerations would be applicable to RIO judgements in the group judgements round.

First, consider the single event of getting "heads" on a single toss of an ordinary coin, where there is no reason to suspect it is biased in any way. I have no trouble in giving the probability 0.5 – i.e. I give a 50% probability to the event happening and 50% to it not happening. Now compare this with the event of a drawing-pin (also known as a thumbtack) falling in the position shown here when tossed onto a table. I feel that the drawing-pin is about equally likely to land this way ("pin up") or on its side ("pin down"), so I give this event a probability of 0.5 too, but this somehow feels like a less confident judgement than the 0.5 probability for tossing "heads".

This kind of reasoning leads many people to feel that I ought to express more uncertainty about P(pin up) = 0.5 than about P(heads) = 0.5.

[2] These examples concern events that are repeatable, but then we will move on to discuss non-repeatable events, like the event of a terrorist flying an aircraft into a nuclear reactor at some point during its operating life.

Many ways have been proposed to address this perceived problem, including "levels of confidence" for probability judgements[3], probability distributions for probabilities and "imprecise probabilities"[4].

However, these ideas fail to recognise the simple fact that uncertainty about a single event is completely described by a single probability. For a single toss, both "heads" and "pin up" have the same probability for me; my uncertainty about them is exactly the same. But my uncertainty about the corresponding frequencies F_H (for "heads") and F_U (for "pin up") are quite different. In a population of a potentially infinite number of coin tosses, I have very little uncertainty about the frequency of "heads". I would expect to see "heads" on very close to 50% of tosses. It might not be exactly 0.5 because a coin is not completely symmetric, but I would be extremely surprised if "heads" occurred on more than 60% or fewer than 40% of tosses in the long run. My probability distribution for F_H is narrowly concentrated around 0.5.

In contrast, I would not be at all surprised to see "pin up" occurring in more than 60% or fewer than 40% of drawing-pin tosses. In fact, I would not be very surprised to see F_U being more than 0.75 or less than 0.25. I have much more uncertainty about F_U than about F_H, and this would be shown in a much more widely spread probability distribution for F_U.

How probabilities change

In general, the sense of confidence, or lack of confidence, in a probability judgement is at least in part a matter of how robust, or sensitive, it might be if new information were to become available. For repeatable quantities, we might imagine observing the values in a sample of other individuals in the population. Thus, if I had observed "pin up" on each of three previous tosses of the drawing-pin, then my probability for "pin up" on the particular toss in question would be appreciably more than 0.5, and this is related to my feeling that the underlying frequency F_U is very uncertain. Whereas if I had previously observed "heads" on three tosses of the coin my probability for "heads" for a single toss would not change. I would attribute the three "heads" to random chance, because I am very confident that the frequency F_H will be close to 0.5.

Even if the quantity of interest is the outcome of a repeatable event E for a single, specified individual, and we have no intrinsic interest in the underlying frequency, it may still be useful to elicit the distribution of the frequency F_E in order to express the robustness of the elicited single probability P(E). This also permits a check on the elicited judgements

[3] The Intergovernmental Panel on Climate Change (IPCC) assigns loosely defined 'confidence' levels to probabilities – see http://www.environment.gov.au/climate-change/publications/fact-sheet-confidence-likelihood.

[4] See the Wikipedia article: https://en.wikipedia.org/wiki/Imprecise_probability.

because it can be shown[5] that P(E) should equal the expected value of the distribution of F_E.

The general case

If the quantity of interest is not repeatable, as in the case of the event W in Example 1 or the quantity N in Example 2, it can still be valuable to consider how your probability judgements would change if more information became available.

Suppose that your probability for an event E would change if you knew the value of some quantity Y. Y might be another event, taking only two values, in which case you would have a probability for E if Y occurs and a different probability if it doesn't, or Y might take several or many possible values. Using standard probability notation, we denote your probability for E if Y takes some value y by $\mathrm{P}(E \mid Y = y)$.

The amount by which $\mathrm{P}(E \mid Y = y)$ changes with the possible values y of Y is one indicator of the confidence with which you can assign a probability to E without knowing Y. But the other important factor is how likely different values of Y might be, so you also need to elicit a probability distribution for Y. (In the special case when Y is an event, this is just the probability of it occurring, together with one minus that probability for it not occurring.)

As in the discussion of frequencies, we can show that P(E) must equal the expected value of your conditional probability $\mathrm{P}(E \mid Y)$, and similarly this provides a way to check value for P(E) that has been elicited through the SHELF 3 Discrete template.

More importantly, however, this approach often provides a way to elicit P(E) with increased confidence. When a direct elicitation of the probability of an event (or the set of probabilities for a discrete quantity) is difficult and leaves a feeling of lack of confidence in the judgement(s), experts may feel much more confident in making judgements conditional on an *extension variable* Y. Consider the event W of winning the contract in Example 1. An expert may feel much more confident in assigning a probability to W if the value of N, the number of other bidders, in Example 2 were known. Then N becomes the extension variable used to elicit a probability for W.

SHELF version 4 includes a new template, "SHELF 3 Extension" for this method (and for the more general case where we elicit a distribution for a quantity X, rather than for an event E).

There is much more discussion of the use of the SHELF 3 Extension template in the SHELF advice document "Extension", and this kind of elaboration is also mentioned in the "Definitions" document.

---

[5] This result is known as de Finetti's theorem.