Eliciting Multivariate Probability Distributions

Alireza Daneshkhah and Jeremy Oakley

September 10, 2010

1 Introduction

In this chapter we consider how to elicit a multivariate distribution to represent an expert's uncertainty about a vector variable $\theta = (\theta_1, \dots, \theta_d)$. This is an extension of the univariate case discussed in the previous chapter, which we assume the reader has studied. We again have three individuals in the elicitation process: a decision-maker who requires the distribution for θ, a (female) expert, whose beliefs about θ the decision-maker wishes to elicit, and a (male) facilitator who will conduct the expert elicitation session.

Multivariate elicitation schemes typically involve eliciting univariate distributions as part of the process, and so all the methods discussed and issues raised in the previous chapter are relevant here. Modifications and extensions of univariate methods are needed, and we start with the overall process of elicitation. Adapting the model from the previous chapter, we have the following.

1. The decision-maker identifies a need for expert judgement, and recruits an expert and facilitator.
2. The facilitator trains the expert in the process of univariate elicitation, so that the expert is comfortable with the idea of representing uncertainty with subjective probability distributions.
3. The facilitator discusses the concept of dependence with the expert, and helps her identify which uncertain quantities should be treated as dependent and which should be treated as independent
4. The facilitator gives the expert further training in multivariate elicitation techniques, so she is aware of the options for constructing joint probability distributions.

5. The facilitator and expert discuss the available evidence, and decide the precise role of elicitation. This includes the "structuring" stage, where it is decided precisely what variables to elicit, and how the joint distribution for θ will be constructed.

6. The facilitator elicits a probability distribution for θ from the expert.

1.1 Outline of the remaining chapter

In the next section, we discuss the concept of dependence between random variables. The presence or absence of dependence determines whether multivariate elicitation techniques are required. We then discuss various multivariate elicitation methods: using simple transformations to obtain independence; constructing a prior hierarchically under a judgement of exchangeability; direct elicitation of multivariate normal distributions; copula methods. We finish with the discussion of one specific multivariate elicitation problem: eliciting beliefs about the parameters of a normal linear model.

2 Dependence

The definition of independence is often given as follows: random events A and B are independent if $P(A, B) = P(A)P(B)$, i.e. their joint probability is the product of the two marginal probabilities. If $P(A, B) \neq P(A)P(B)$ then we say that A and B are dependent. More intuitive for our purposes is the equivalent definition:

> Random variables A and B are independent if $P(A|B) = P(A)$.

In subjective probability theory, we read this to mean that if the expert judges A and B to be independent, she would not change her beliefs about A given (new) knowledge of B, and vice versa. The events A and B do not have to have any form of physical or causal relationship for her to judge them to be dependent. Dependence is a property of *her beliefs* about A and B.

In the elicitation, a simple thought experiment for the expert to try is the following. Supposing she were to estimate two uncertain variables θ_1 and θ_2 by m_1 and m_2, where m_1 and m_2 are the medians of her marginal distributions for θ_1 and θ_2 respectively. By definition, she should not judge the event $\{\theta_i < m_i\}$ to be more likely or less likely than the event $\{\theta_i > m_i\}$, i.e. underestimating θ_i with m_i should not be more or less likely that overestimating θ_i, for $i = 1, 2$. Now suppose she were to learn that she had underestimated θ_1. Would this make her think that she had underestimated θ_2 also? If she now judges that the events $\{\theta_2 < m_2\}$ and $\{\theta_2 > m_2\}$ are not equally likely (with either one having a higher probability), then she must judge θ_1 and θ_2 to be dependent. Again, the point here is not whether the event $\{\theta_1 > m_1\}$ *causes* $\{\theta_2 > m_2\}$, but whether learning $\{\theta_1 > m_1\}$ changes the expert's beliefs about the event $\{\theta_2 > m_2\}$.

The importance of identifying dependence and using multivariate elicitation will depend on how the elicited distribution is to be used. Suppose for example that the decision maker is interested in the value of the function $h(\theta_1, \theta_2) = \theta_1 + \theta_2$, in addition to θ_1 and θ_2 individually. If we consider the expert's uncertainty about $h(\theta_1, \theta_2)$ as described by her variance, we have $Var\{h(\theta_1, \theta_2)\} = Var(\theta_1) + Var(\theta_2) + 2Cov(\theta_1, \theta_2)$. If the facilitator ignores dependence and supposes that $Var\{h(\theta_1, \theta_2)\} = Var(\theta_1) + Var(\theta_2)$ (where $Var(\theta_1)$ and $Var(\theta_2)$ are obtained through eliciting the expert's marginal distributions of θ_1 and θ_2), he may be understating or overstating her uncertainty depending on her covariance between θ_1 and θ_2.

We now state an important objective in multivariate elicitation that is unique to the multivariate case. As well as ensuring that the expert's uncertainty about each individual θ_i is faithfully represented, the facilitator must ensure that the expert's beliefs about *joint* events such as $\{\theta_i < a, \theta_j < b\}$ is faithfully represented. The idea of feedback, as discussed in the previous chapter, plays an important role here. Eliciting joint beliefs is difficult, and whatever method is used, the facilitator should explore the joint elicited distribution carefully with the expert. For any joint events of importance, the expert should be satisfied that her joint beliefs have been described appropriately.

3 Transformation methods

If an expert judges θ_1 and θ_2 to be dependent, it may be possible to find transformations $\phi_1 = g_1(\theta_1, \theta_2)$ and $\phi_2 = g_2(\theta_1, \theta_2)$ such that she judges ϕ_1 and ϕ_2 to be independent and is willing to express beliefs about ϕ_1 and ϕ_2. Dependence between θ_1 and θ_2 is induced through deriving their joint distribution from the distributions of ϕ_1 and ϕ_2. An illustration of this approach is given in O'Hagan (1998).

3.1 Mean and difference

As a simple example, suppose θ_1 and θ_2 are the heights (in cm) of two adult brothers. The expert might expect θ_1 and θ_2 to be similar, so that she would change her beliefs about θ_1 if she observed θ_2. For example, suppose she has judged that $\theta_i \sim N(175, 7^2)$ for $i = 1, 2$. If she then learns that $\theta_1 = 190$cm, she might expect θ_2 to be closer to 190cm than to 160cm.

We could instead consider the expert's beliefs about $\phi_1 = (\theta_1 + \theta_2)/2$, and $\phi_2 = \theta_1 - \theta_2$, i.e., the brothers' mean height and the difference between their two heights, and judge that these two quantities are independent. For example, if we learn that the brothers differ in height by 1cm, we may not feel that we have learnt anything new about their meanheight.

Distributions for ϕ_1 and ϕ_2 can be elicited using univariate techniques, as described in the previous chapter. We then transform back to obtain the joint distribution of θ_1, θ_2. For example, if $\phi_1 \sim N(m_1, v_1)$ and $\phi_2 \sim N(m_2, v_2)$ then

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} \frac{2m_1+m_2}{2} \\ \frac{2m_1-m_2}{2} \end{pmatrix}, \begin{pmatrix} v_1 + \frac{v_2}{4} & v_1 - \frac{v_2}{4} \\ v_1 - \frac{v_2}{4} & v_1 + \frac{v_2}{4} \end{pmatrix} \right\}$$

3.2 Baseline and relative risk

This parameterisation is common in medical applications. Suppose we have two treatments for the same illness, with θ_1 and θ_2 the population proportion of patients who would respond to each treatment. Again, knowledge of θ_1 would be informative for θ_2. For example, if the expert learns that $\theta_1 = 0.1$, this tells her something about how difficult the illness is to cure, and she might expect θ_2 to be closer to 0.1 than to 0.9.

Using the baseline and relative risk parameterisation, we elicit beliefs about θ_1 and $\phi = \theta_2/\theta_1$, that is, we consider how much more (or less) effective the second treatment could be, compared to the first. We might, for example, elicit a beta distribution to represent uncertainty about θ_1, and a lognormal distribution to represent uncertainty about ϕ. Though this would not give us a standard parametric distribution for θ_2, it would be straightforward to simulate from the distribution of θ_2. Some illustrations of this approach are given in Nixon et al. (2009), Stevenson et al. (2009a) and Stevenson et al. (2009b).

As another example, consider forecasting gross domestic products (GDP) of Eurozone countries. An expert might first consider the average GDP of Eurozone countries, represented by the variable θ_1. She could then use θ_1 as a baseline to forecast the GDP θ_2 of a specific European country by means of the ratio $\phi = \theta_2/\theta_1$. The benefits of baseline and relative risk estimates are also stressed in the Chapter by Riccardo Rebonato about financial stress testing in this book.

4 Exchangeability, conditional independence, and hierarchical priors

In some cases it may be possible to simplify the elicitation problem by constructing a prior *hierarchically*. This can be useful when we have a large number of uncertain quantities. Consider the following example. Suppose $\theta_1, \ldots, \theta_d$ are the race times of a male 10Km track runner on d separate occasions, and we wish to elicit an expert's joint distribution $f(\theta_1, \ldots, \theta_d)$. She judges that $\theta_1, \ldots, \theta_d$ are not independent, as θ_i is informative for θ_j. For example, if she learnt that $\theta_1 = 28$ minutes, this would suggest that the runner is of Olympic standard, and she would expect similarly fast times for $\theta_2, \ldots, \theta_d$.

The facilitator can construct a hierarchical prior if the expert judges $\theta_1, \ldots, \theta_d$ to be *exchangeable*. Informally, she would judge $\theta_1, \ldots, \theta_d$ to be exchangeable if she had no reason to make different assessments about any two θ_i and θ_j. If, for example, she expected the runner's times to improve on each occasion, then exchangeability would not hold. The formal requirement for exchangeability is that for all values $a_1, \ldots, a_d$, and all $1 \leq i_1 < i_2 < \ldots < i_d$, the following holds:

$$P(\theta_1 < a_1 \cap \theta_2 < a_2 \cap \ldots \theta_k < a_d) = P(\theta_{i_1} < a_1 \cap \theta_{i_2} < a_2 \cap \ldots \theta_{i_d} < a_d),$$

i.e. the indices of $\theta_1, \ldots, \theta_d$ can be permuted without changing the probabilities (De Finetti, 1975, section 11.4).

Under exchangeability, each θ_i has the same marginal distribution, and so it would not be necessary to elicit each of $f(\theta_1), \ldots, f(\theta_d)$. The facilitator can construct the prior hierarchically: the marginal prior for each θ_i is specified in terms of some unknown hyperparameters, and uncertainty about these hyperparameters are represented with hyperprior distributions. For example, the prior could be of the form

$$\begin{aligned}
\theta_i|\mu, \sigma^2 &\sim N(\mu, \sigma^2) \text{ for } i = 1, \ldots, d, \\
\mu &\sim N(m, v), \\
\sigma^{-2} &\sim Gamma(a, b),
\end{aligned}$$

with values elicited for the hyperparameters m, v, a and b.

The judgement here is that $\theta_1, \ldots, \theta_d$ are independent *conditional* on μ and σ^2. Suppose the expert is told that the runner's average race time is 30 minutes, and that the variance of his race times is 100 seconds. Assuming his general level of performance does not change over the period in question, the expert might judge that learning θ_i =31:10 in one race would not change her beliefs about θ_j, *if she already knew* the mean and variance of his running times. Unconditional on μ and σ^2, the quantities $\theta_1, \ldots, \theta_d$ are still dependent; θ_i is informative for θ_j, because knowledge of θ_i would change the expert's beliefs about μ and σ^2, and changing her beliefs about μ and σ^2 would change her beliefs about θ_j.

In this example, μ and σ^2 have intuitive interpretations: μ is the runner's average time, and σ^2 describes variability between running times on different occasions. It make sense to think of these parameters as uncertain quantities (in that they really would be unknown to the expert), which is a good indication that the hierarchical prior is a helpful way to think about the elicitation problem.

Such hierarchical models are already wide spread in financial modelling. For example, in the classic Basel II one-factor model for credit risk, default of an individual borrower i ($i = 1, ..., d$) occurs when its ability-to pay-variable θ_i falls below a certain threshold (the default point). Now, by empirical evidence it is clear that different borrowers' defaults are dependent, however, similar to our track runner example above, it is assumed that the $\theta_1, ..., \theta_d$ are independent conditional on a single macroeconomic factor Y.

4.1 Eliciting hyperpriors in hierarchical models

In general, it is considered good practice to ask for judgements about *observable quantities* rather than to ask for judgements about parameters in statistical models (Kadane and Wolfson, 1998), though the latter is acceptable if the expert has a good intuitive understanding of what a particular model parameter represents. To elicit the hyperpriors, the facilitator should therefore ask for suitable judgements about $\theta_1, \ldots, \theta_d$, and infer hyperparameters from these judgements. This is known as predictive elicitation (see for example Percy, 2004), and we illustrate this approach later in this chapter in elicitation for linear models.

5 Eliciting bivariate normal distributions

When methods based on transformation or hierarchical modelling are not helpful, the facilitator may need to elicit a joint distribution directly. Choices of multivariate distribution are limited, with the most obvious being the multivariate normal distribution. We concentrate on the bivariate normal case before considering extensions to higher dimensions.

Continuing the running theme, suppose the two uncertain quantities are the men's world records for the 5Km and 10Km track races, denoted by θ_1 and θ_2, at some specified time in the future. We suppose that the expert judges θ_1 and θ_2 to be dependent, for the reason that learning that she had underestimated θ_1 would make her believe it more likely that she had underestimated θ_2 also. Clearly θ_1 and θ_2 are not exchangeable, and so the facilitator chooses to elicit a multivariate distribution directly, in this case, the bivariate normal:

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \begin{pmatrix} v_1 & v_{12} \\ v_{12} & v_2 \end{pmatrix} \right\} \tag{1}$$

The first step is to elicit marginal distributions for θ_1 and θ_2, which can be done using the univariate procedures described in the previous chapter. This will give m_1, m_2, v_1 and v_2, since $\theta_1 \sim N(m_1, v_1)$ and $\theta_2 \sim N(m_2, v_2)$. This leaves the covariance parameter v_{12} to be elicited, and there are various options for doing so.

5.1 Eliciting joint probabilities

The facilitator can directly elicit joint probabilities of the form

$$P(\theta_1 \leq a, \theta_2 \leq b) \quad \text{or} \quad P(a_1 < \theta_1 \leq a_2, b_1 < \theta_2 \leq y_b)$$

Moala and O'Hagan (2010) ask for such judgements in their nonparametric elicitation method. For an arbitrary choice of a_1, a_2, b_1, b_2, we cannot always identify a unique covariance parameter v_{12}, and so it may be necessary to elicit several joint probabilities, and find (numerically) the value v_{12} that gives the best fit to the elicited probabilities.

Another option, used in Fackler (1991), is to ask for a quadrant probability:

$$p_q = P(\theta_1 > m_1, \theta_2 > m_2),$$

i.e. the probability that both θ_1 and θ_2 are above their expected values. This is more convenient mathematically, as we can identify v_{12} directly:

$$v_{12} = \sin\{2\pi(p_q - 0.25)\}\sqrt{v_1 v_2}.$$

5.2 Eliciting conditional probabilities

Instead of asking for joint probabilities, the facilitator can ask for conditional probabilities. Given the bivariate normal distribution in (1), the conditional distribution $\theta_1 | \theta_2 = \theta_2^*$ is also normal:

$$\theta_1 | \theta_2 \sim N\left\{ m_1 + v_{12} v_{22}^{-1}(\theta_2^* - m_2), v_{12}^2 v_{22}^{-1} \right\}$$

If the facilitator has already obtained m_1, m_2, v_1 and v_2 by eliciting the two marginal distributions, he can then elicit beliefs about $\theta_1 | \theta_2$ by suggesting a hypothetical value θ_2^* and using univariate elicitation methods. In principle, eliciting the expert's median value of θ_1 conditional on θ_2 would be sufficient (provided the hypothetical value θ_2^* is not chosen to be m_2). Denoting the elicited conditional median by $m_{1|2}$, we have

$$v_{12} = \frac{m_{1|2} - m_1}{v_{22}^{-1}(\theta_2^* - m_2)}.$$

We recommend eliciting additional judgements (for example the inter-quartile range, as in the bisection method), and checking for consistency. If the facilitator elicits the inter-quartile range, denoted by *iqr*, we would then have

$$v_{12} = \frac{iqr\sqrt{v22}}{1.349}.$$

The facilitator can also repeat the exercise for different hypothetical values (or for θ_2 conditional on θ_1), both to check for consistency and to offer feedback.

This approach illustrates a technique used in various multivariate elicitation schemes (including elicitation for linear models). The expert is given some hypothetical data, and asked to make judgements conditional on this data. The facilitator then infers the required parameter values from these conditional judgements, on the assumption that the the expert has followed the laws of probability precisely when revising her beliefs. Of course, we don't believe that this is what the expert *actually* does, but the assumption gives the facilitator a means of constructing the distribution, and he can (and should) always use feedback to see if the elicited distribution is acceptable to the expert. Note there is a trade-off to be made between the ease of considering univariate probabilities, with the difficulty of mentally conditioning on hypothetical data. Unfortunately, we are not aware of any empirical evidence suggesting which elicitation method is more reliable.

5.3 Extensions to higher dimensions

In principle, the previous methods can be used for multivariate normal distributions with $d = 3$ or higher. However, there is no guarantee that the elicited variance matrix will be positive definite. Problems can occur, for example, if the experts judges θ_1 to have strong positive correlation with both θ_2 and θ_3, but θ_2 and θ_3 to have strong negative correlation. The facilitator will need to monitor the pairwise correlations carefully, and give feedback and advice if the variance matrix is not valid.

6 Eliciting multivariate distributions with copulas

The copula method is a general approach for constructing multivariate distributions. In this method, the facilitator must elicit a marginal distribution for each variable, and correlations between each pair of variables. An advantage of this method is its flexibility in terms of the choice of marginal distributions: any parametric distribution can be specified for each marginal. The drawback is that eliciting correlations can be difficult. We give a short review here, but further details can be found in Clemen and Reilly (1999) and Kurowicka and Cooke (2006).

The basis for the copula method is the following theorem, due to Sklar (1959):

> A joint cumulative distribution function $F(x_1, \ldots, x_n)$ for $X_1, \ldots, X_n$ with respective marginal distribution functions $F_1(x_1), \ldots, F_n(x_n)$ can be written as a function of its marginals:
>
> $$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$
>
> where $C(\cdot)$ is a joint distribution with uniform marginals called the copula. If each $F_i(\cdot)$ is a continuous distribution function then $C(\cdot)$ is unique, and if each $F_i(\cdot)$ is discrete, then $C(\cdot)$ is unique on $Ran(F_1) \times Ran(F_n)$, where $Ran(F_i)$ is the range of $F_i(\cdot)$.

If each F_i is differentiable, we can rewrite the joint density

$$f(x_1, \ldots, x_n) = f_1(x_1) \times \ldots \times f_n(x_n) \times c(F_1(x_1), \ldots, F_n(x_n)), \tag{2}$$

where $f_i(\cdot)$ is the density corresponding to $F_i(\cdot)$ and $c = \partial^n C/(\partial F_1 \ldots \partial F_n)$, which is called the copula density.

From equation (2) we see that we can elicit a joint density by eliciting each marginal density and the copula density. The marginal densities can be elicited using univariate methods from the previous chapter, and so we now consider how to elicit the copula density. The facilitator must choose the form of the copula density, and one option, which we discuss here, is the multivariate Gaussian copula. An alternative is the minimum information copula, which has minimal information with respect to the uniform (independent) copula amongst all those copulae with a given Spearman rank correlation (see Bedford and Meeuwissen, 1997).

The multivariate Gaussian copula is parameterised by a (product-moment) correlation matrix R and has the general form

$$c_N[\Phi(y_1), \ldots, \Phi(y_n) \mid R] = \exp\{\mathbf{y}^T(R^{-1} - I)\mathbf{y}/2\}/|R|^{1/2}$$

where $\mathbf{y} = (y_1, \ldots, y_n)^T$, I is the $n \times n$ identity matrix, and $\Phi(\cdot)$ denote the univariate standard normal distribution. The use of this copula in (2), with $y_i = \Phi^{-1}[F_i(x_i)]$ for $i = 1, \ldots n$, gives

$$\begin{aligned} f(x_1, \ldots, x_n \mid R) \quad = \quad & f_1(x_1) \times \ldots \times f_n(x_n) \times |R|^{-\frac{1}{2}} \times \exp\{-(\Phi^{-1}[F_1(x_1)], \ldots, \Phi^{-1}[F_n(x_n)]) \\ & \times(R^{-1} - I) \times (\Phi^{-1}[F_1(x_1)], \ldots, \Phi^{-1}[F_n(x_n)])^T/2\}. \end{aligned}$$

Having elicited each marginal, this leaves the problem of eliciting the matrix R. Putting aside the issue of how one might assess a correlation, the facilitator cannot ask for Pearson's product-moment correlations directly, as the expert's product-moment correlation between X_i and X_j does not necessarily equate to element i, j of R, depending on the choice of marginal distributions. Instead, the expert must assess rank-order correlations (such as Spearman's ρ or Kendall's τ). If R^* is an elicited matrix of rank-order correlations, then we obtain R from R^* using $R_{ij} = \sin(\pi R^*_{ij}/2)$ if Kendall's τ has been elicited, and $R_{ij} = 2\sin(\pi R^*_{ij}/6)$ if Spearman's ρ has been elicited. A problem here is that there is no guarantee that the resulting matrix R is positive definite, and so this must be checked by the facilitator, with the facilitator and expert agreeing adjustments if necessary. The chapter by Böcker, Crimmi, and Fink in this book gives an example where this technique is used to elicit the Gaussian copula parameter between a bank's different risk types.

6.1 Eliciting correlations

Various methods for eliciting correlations are suggested in Clemen and Reilly (1999), and a more substantial investigation is reported in Clemen et al. (2000). However, the methods described by these authors are typically intended for *estimating* correlations in populations rather than for eliciting dependence in beliefs about two fixed, uncertain quantities. Methods for the former task are not necessarily suitable for the latter (a point which the authors discuss clearly). The distinction is important, and so we illustrate it with the following example.

Suppose there is a population of adults, with X_i and Y_i the height and weight of the i-th member of the population respectively. We define μ_X and σ_X^2 to be the population mean and variance of the heights, and μ_Y and σ_Y^2 to be the population mean and variance of the weights. This population has a true product-moment correlation between height and weight, defined by

$$\rho = \frac{E\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y},$$

and we could ask an expert to estimate ρ (or even consider uncertainty about ρ). However, in the elicitation context, we are typically interested in uncertainty in fixed quantities, for example μ_X and μ_Y. An expert may judge μ_X and μ_Y to be dependent, in that she would revise her beliefs about μ_X given new knowledge of μ_Y (and vice versa), but there is no *true* correlation coefficient for these two parameters.

One method given in Clemen and Reilly (1999) that can be used (with some modification to their wording) is the conditional fractile method. Given marginal distributions for uncertain quantities θ_1 and θ_2, Spearman's correlation can be assessed by asking questions such as following:

"Suppose you were to learn that θ_1 took the value a_1, the 100α percentile of its distribution. What would you estimate θ_2 to be?"

Suitable values of a_1 and α are chosen using the marginal distribution of θ_1. Given the estimate of θ_2, Spearman's correlation is determined based on a nonparametric regression relationship

$$E\{F(\theta_2)|\theta_1\} = \rho_{\theta_1\theta_2}\{F(\theta_2) - 0.5\} + 0.5.$$

Clemen and Reilly (1999) suggest asking for several estimates and choosing $\rho_{\theta_1\theta_2}$ using a least squares fit.

Given the difficulty in making these sorts of assessments, a pragmatic option is for the facilitator to use trial and error, with appropriate feedback. Having established whether a particular correlation should be positive or negative, the facilitator can choose elements of R directly, and report joint summaries back to the expert. These could be quadrant probabilities $P(\theta_1 > m_1, \theta_2 > m_2)$, where m_i is the median of the expert's distribution for θ_i, or even scatterplots showing samples from the joint distribution of θ_1, θ_2. The elements of R can then be adjusted until the expert is satisfied that her joint beliefs are being represented appropriately.

An alternative approach is to elicit conditional correlations, where the expert considers dependence between X_i and X_j conditional on X_k. This can avoid the difficulty of ensuring positive definiteness mentioned earlier. This forms the basis of the vine approach presented in Bedford and Cooke (2002), and further discussion of the elicitation of conditional rank correlations is given in Morales et al. (2008). Another alternative is given in Bedford (2006), in which correlations are elicited through judgements about observable quantities. Here the expert is given feedback regarding what values can be specified to ensure coherence. Further details and financial applications are given in Daneshkhah and Bedford (2010a,b) and Lewandowski (2008).

7 Eliciting beliefs about the parameters of a normal linear model: the method of Kadane et al (1980).

We now consider how to elicit beliefs about the parameters in the normal linear model. We review in detail the method of Kadane, Dickey, Winkler, Smith and Peters (1980) (hereafter KDWSP), which was developed further in Kadane and Wolfson (1998). One alternative approach, which we do not consider here, is given in Garthwaite and Dickey (1988).

The method is complex and illustrates the difficulties of multivariate elicitation, both with regards to the theory required for constructing valid multivariate distributions, and with regards to the judgements required by the expert. In this chapter we describe the method and give a numerical example, but the reader should consult KDWSP for the underlying theory.

7.1 Example: key risk indicators for operational risk

We illustrate the method with the following example borrowed from operational risk, which can be defined as the risk of losses resulting from inadequate or failed processes, systems, or external events. In operational risk management, an important concept is that of key risk indicators (KRIs). Basically, KRI are used to systematically explain the amount and severity of operational loss events by a couple of significant key factors (further details on KRIs can be found in Davies et al., 2006). Note also that this example is used in the chapter by Ioannis Ntzoufras about Bayesian analysis of the normal regression model in this book. R code for implementing this example is available at `http://www.jeremy-oakley.staff.shef.ac.uk/LMelicitation.R`

We suppose that risk manager wishes to elicit an expert's beliefs about the relationship between *operational losses* y and five explanatory variables, *cycle time and timeliness of transactions* (x_1), *transaction volume* (x_2) , *hiring and training costs* (x_3), *customer salification index* (x_4) and *IT Network downtime* (x_5). The facilitator and expert discuss the nature of this relationship, and agree on representing it using a linear model:

$$Y = \mathbf{x}'\beta + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2), \tag{3}$$

with $\mathbf{x}' = (1, x_1, \ldots, x_{r-1})$, and $r = 6$ here.

The uncertain parameters are the regression coefficients β and the error variance σ^2, and so the facilitator needs to elicit the expert's prior beliefs about β and σ^2.

As discussed earlier in elicitation for hierarchical models, we typically do not expect the expert to be willing to make judgements about model parameters directly, even though she has subject matter expertise about the relationship between Y and $\mathbf{x}$. It would be hard to make judgements about σ^2 directly, and the individual elements of β could become difficult to interpret if $\mathbf{x}$ contained nonlinear functions of the independent variables. KDWSP ask the expert to make judgements about *observable quantities*, and then infer distributions for the model parameters from these judgements. The observable quantity here is Y, and we suppose that the expert is willing to make judgements about Y given $\mathbf{x}$.

7.2 Beliefs about model parameters

If the expert is unwilling to make direct judgements about β and σ^2, what do we mean by her "beliefs about β and σ^2" in this case? Let $X = (\mathbf{x}_1, \ldots, \mathbf{x}_m)'$ be an $m \times r$ matrix corresponding to m different choices of the explanatory variables, with $\mathbf{Y}$ defined as $X\beta + \varepsilon$, where ε is the corresponding vector of independent normally distributed errors. Any choice of distribution $p(\beta, \sigma^2)$ implies a predictive distribution for $\mathbf{Y}$ given X, through

$$p(\mathbf{Y}|X) = \int\int p(\mathbf{Y}|X, \beta, \sigma^2) p(\beta, \sigma^2) d\beta d\sigma^2, \tag{4}$$

where $\mathbf{Y}|X, \beta, \sigma^2$ has the multivariate normal distribution $N_m(X\beta, \sigma^2 I)$ distribution. Hence, when we refer to "the expert's beliefs about β and σ^2", we imagine that a particular choice of $p(\beta, \sigma^2)$ will result in (4) matching the expert's beliefs about $\mathbf{Y}|X$, so that this choice of $p(\beta, \sigma^2)$ describes the expert's (implicit) beliefs about the parameters.

Of course, no expert would mentally evaluate the integral (4) to make her judgements about $\mathbf{Y}$, and almost certainly that there will not be any distribution $p(\beta, \sigma^2)$ that gives a perfect match (for large enough m). Neverthless, being pragmatic, considering (4) gives us a framework for eliciting a candidate prior for (β, σ^2). We can then use feedback (by reporting summaries from the distribution of Y for different $\mathbf{x}$) to see if the elicited prior is an acceptable representation of the expert's beliefs.

In theory, the facilitator could elicit the predictive distribution of the observables $\mathbf{Y}$, and then construct a prior $p(\beta, \sigma^2)$ so that (4) matches this elicited distribution as closely as possible. Doing so would be difficult, both in the requirement for multivariate assessments about $\mathbf{Y}$, and in identifying a suitable functional form for $p(\beta, \sigma^2)$. KDWSP show that at the cost of imposing a conjugate prior on the expert's beliefs, it is possible to identify the parameters of the prior using a series of univariate elicitation tasks. Given the difficulty of the elicitation task for the expert (regardless of the method used), and likely presence of some imprecision in her assessments, we think the imposition of a conjugate prior will almost always be acceptable in practice.

7.3 A conjugate prior for β and σ^2

The conjugate joint prior distribution of (β, σ) is the multivariate normal inverse chi-squared distribution. Using KDWSP's notation [1]:

$$p(\beta|\sigma^2) \sim N\left(\mathbf{b}, \frac{\sigma^2 R^{-1}}{\delta + r}\right), \tag{5}$$

[1] Note that often the prior for σ^2 is expressed in terms of an inverse gamma distribution which gives an equivalent prior specification.

$\mathbf{x}$	x_1	x_2	x_3	x_4	x_5	$y_{0.5}$	$y_{0.75}$	$y_{0.9}$
$\mathbf{x}_1$	20000	250	10	600	4	385	410	430
$\mathbf{x}_2$	70000	250	10	200	8	200	235	265
$\mathbf{x}_3$	20000	500	10	200	8	575	605	630
$\mathbf{x}_4$	70000	500	10	600	4	260	295	330
$\mathbf{x}_5$	20000	250	20	600	8	700	735	770
$\mathbf{x}_6$	70000	250	20	200	4	390	430	470
$\mathbf{x}_7$	20000	500	20	200	4	765	800	840
$\mathbf{x}_8$	70000	500	20	600	8	575	620	655

Table 1: Design points, and the expert's elicited medians, 75th and 90th percentiles

$$\frac{\omega\delta}{\sigma^2} \sim \chi^2_\delta, \tag{6}$$

where R is an $r \times r$ positive definite symmetric matrix and δ, ω and σ^2 are all positive. The facilitator has to obtain the hyperparameters $\mathbf{b}, R, \omega$ and δ. Note that a conjugate prior is not being used specifically so we could easily derive a posterior given new data; there may not be any new data. The reason for using a conjugate prior is that (4) can be evaluated analytically (to a give a multivariate t density), so that we can work back from judgements about the observable Y to implied beliefs about the parameters.

7.4 Eliciting b

The facilitator and expert first choose the design matrix $X = (\mathbf{x}_1, \ldots, \mathbf{x}_m)'$ referring to m different realisations of the KRI. Here, we use a fractional factorial design with $m = 8$, allowing each independent variable to take a low or high value. It is possible that combinations of independent variable values may not be plausible which would make it hard for the expert to consider her beliefs about Y, but we do not consider this here. Kadane and Wolfson (1998) suggest an algorithm for choosing X, in which the expert can 'reject' implausible values of $\mathbf{x}$. A choice of X is given in columns 2-6 of Table 1.

For each $\mathbf{x}_i$, the facilitator elicits the expert's median, 75th percentile and 90th percentile of her distribution for $Y_i = \mathbf{x}_i'\beta + \varepsilon_i$. For the i-th design point, we denote these three judgements by $y_{i,0.5}$, $y_{i,0.75}$ and $y_{i,0.9}$. The bisection method discussed in the previous chapter can be used to elicit the median and 75th percentile. The expert would specify her 90th percentile directly. Illustrative values are given in Table 1.

To obtain $\mathbf{b}$, the facilitator treats the elicited medians as observations of $Y_1, \ldots, Y_m$, and sets $\mathbf{b}$ equal to the corresponding least squares estimate:

$$\mathbf{b} = (X'X)^{-1}X\mathbf{y}_{0.5}, \tag{7}$$

with $\mathbf{y}_{0.5} = (y_{1,0.5}, \ldots, y_{m,0.5})'$. The facilitator can now give the expert some feedback, by comparing $\mathbf{y}_{0.5}$ with $X\mathbf{b}$, with attention drawn to any large differences. In the example facilitator would obtain

$$\mathbf{b} = (48.75, -0.005, 0.5, 25.25, -0.00625, 15.625)'.$$

7.4.1 Choice of X: further discussion

An alternative to the fractional factorial design used in Table 1 would be to choose $\mathbf{x}_1$ with each independent variable set at some central value, and then vary each independent variable one at a time (e.g. $\mathbf{x}_1 = (1, 45000, 350, 20, 400, 6)$, $\mathbf{x}_2 = (1, 70000, 350, 20, 400, 6)$, $\mathbf{x}_3 = (1, 45000, 500, 20, 400, 6)$ and so on). Arguably, this would be easier for the expert, as she only has to consider the effect of changing a single independent variable when considering beliefs about $Y_2, Y_3, \ldots$. However, she must still consider the effect of all five independent variables when making judgements about Y_1, and if she is able to do so for $\mathbf{x}_1$ then she should be able to do so for any other $\mathbf{x}$. In any case, we recommend that the facilitator explores alternative ways of assessing $\mathbf{b}$ with the expert. For example, he can ask to her to consider the effect on Y of switching x_i from its 'low' setting to its 'high' setting, which could also give an assessment of element $i + 1$ of $\mathbf{b}$.

In the context of linear regression, fractional factorial designs are superior to one-at-a-time designs of the same size, in that they give smaller variances of parameter estimates. While it may be too simplistic to assume that the elicited medians represent 'true values' plus 'independent errors', we still think it is sensible to follow the principles of good experimental design.

7.5 Eliciting δ

The degrees of freedom parameter δ describes the thickness of the tails of the distribution of any Y, and KWDSP estimate this parameter by considering the tail ratios

$$a_i(\mathbf{x}_i) = \frac{y_{i,0.9} - y_{i,0.5}}{y_{i,0.75} - y_{i,0.5}}.$$

As pointed out in KWDSP, these ratios are independent from the spread and the center of the t distribution and thus only depend on δ. Thus, the a_i can be compared with the corresponding tail ratios of the t-distribution for different degrees of freedom. The facilitator chooses δ such that the difference between the average $\sum a_i(\mathbf{x}_i)/m$ and $t_\delta(0.9)/t_\delta(0.75)$ is minimised, where $t_\delta(q)$ is the q-th quantile of the standard t distribution with δ degrees of freedom.

	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$
$\frac{y_{i,0.9}-y_{i,0.5}}{y_{i,0.75}-y_{i,0.5}}$	1.80	1.86	1.83	2.00	2.00	2.00	2.14	2.00

Table 2: Observed tail ratios

We show the elicited tail ratios in Table 2. Again the facilitator can draw the expert's attention to any large discrepancies. Note that $t_\infty(0.9)/t_\infty(0.75) = 1.90$, so that in our example, the expert has given three sets of judgements (for $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_3$) corresponding to a lighter tailed distribution than the normal distribution. The expert may wish to adjust her assessments, but leaving them unchanged, the facilitator estimates δ to be 11.1.

7.6 Eliciting ω and R

Eliciting ω and R is a more complex task. KDWSP make use of what they (and other authors) term the *center* and *spread* of a multivariate t distribution. If $\mathbf{t}_\delta$ is a standard multivariate t vector with δ degrees of freedom, and we define $\mathbf{z} = \mathbf{a}+B\mathbf{t}_\delta$, then $C(\mathbf{z}) = \mathbf{a}$ is the center of $\mathbf{z}$ and $S(\mathbf{z}) = BB'$ is the spread of $\mathbf{z}$. For $\delta > 1$, so that $E(\mathbf{z})$ exists, $C(\mathbf{z}) = E(\mathbf{z})$, and for $\delta > 2$, so that $Var(\mathbf{z})$ exists, $\frac{\delta}{\delta-2}S(\mathbf{z}) = Var(\mathbf{z})$. The parameters ω and R can be elicited by considering centers and spreads, conditional on hypothetical observations.

We first state the formulae required to obtain R, and then describe what questions the facilitator must ask the expert to apply the formulae. The formulae are derived from properties of the multivariate t distribution, with details given in KWDSP.

The matrix R is obtained from the equation

$$R^{-1} = (X_r'X_r)^{-1}X_r'(U_r - \omega I)X_r(X_r'X_r)^{-1}/(\omega/(\delta + r))$$

where X_r is the matrix formed of the first r rows of X, and U_r is an $r \times r$ matrix, constructed iteratively. (We discuss the elicitation of ω later.) Using a set of hypothetical values $y_1^0, \ldots, y_r^0$ for $Y_1, \ldots, Y_r$, and starting with

$$U_1 = S(Y_1) = \left(\frac{y_{1,0.75} - y_{1,0.5}}{t_\delta(0.75)}\right)^2,$$

which follows from the definition of $\mathbf{z}$ and $S(\mathbf{z})$ given earlier, the facilitator must calculate

$$U_{i+1} = \begin{pmatrix} U_i & U_i T_{i+1} \\ T_{i+1}' U_i & S(Y_{i+1}) \end{pmatrix}$$

with

$$T_{i+1} = \begin{pmatrix} y_1^0 - y_{1,0.5} & C(Y_2 \mid y_1^0) - y_{2,0.5} & \cdots & C(Y_i \mid y_1^0) - y_{i,0.5} \\ y_1^0 - y_{1,0.5} & y_2^0 - y_{2,0.5} & \cdots & C(Y_i \mid y_1^0, y_2^0) - y_{i,0.5} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & C(Y_i \mid y_1^0, \ldots, y_{i-1}^0) - y_{i,0.5} \\ y_1^0 - y_{1,0.5} & y_2^0 - y_{2,0.5} & \cdots & y_i^0 - y_{i,0.5} \end{pmatrix}^{-1} .L_{i+1} \quad (8)$$

$$L_{i+1} = \{C(Y_{i+1} \mid y_1^0) - y_{i+1,0.5}, \ldots, C(Y_{i+1} \mid y_1^0, \ldots, y_i^0) - y_{i+1,0.5}\}', \quad (9)$$

and

$$S(Y_{i+1}) = \frac{S(Y_{i+1} \mid y_1^0, \ldots, y_i^0).[1 + i/\delta]}{1 + \frac{1}{\delta}(y_1^0 - y_{1,0.5}, \ldots, y_i^0 - y_{1,0.5})U_i^{-1}(y_1^0 - y_{1,0.5}, \ldots, y_i^0 - y_{1,0.5})'} + T'_{i+1}U_i T_{i+1} \quad (10)$$

Inspecting the definitions above, we see that the facilitator needs to elicit conditional centers and spreads, of the form $C(Y_i \mid y_1^0, \ldots, y_j^0)$ with $j < i$ and $S(Y_{i+1} \mid y_1^0, \ldots, y_i^0)$. We now describe how these terms are elicited.

1. The facilitator chooses a hypothetical value of Y at $\mathbf{x} = \mathbf{x}_1$. We denote this value by y_1^0. Note that y_1^0 must not equal $y_{1,0.5}$, i.e. the hypothetical value must not be what the expert was 'expecting'. In general, the hypothetical values must satisfy the condition
$$y_i^0 \neq C(Y_i | y_1^0, \ldots, y_{i-1}^0).$$

2. The facilitator elicits the expert's median values of $Y_1, \ldots, Y_r$, conditional on the hypothetical value y_1^0, i.e. she supposes that she has observed $Y = y_1^0$ at $\mathbf{x} = \mathbf{x}_1$, and re-assesses her medians. The conditional center $C(Y_i | y_1^0)$ required in (8) and (9) is set equal to the elicited conditional median for Y_i.

3. The facilitator elicits the expert's 75th percentile for Y_2 conditional on y_1^0. Denoting this elicited percentile by $y^*_{2,0.75}$, the conditional spread $S(Y_2|y_1^0)$ required in (10) is set equal to
$$\left(\frac{y^*_{2,0.75} - C(Y_2|y_1^0)}{t_{\delta+1}(0.75)} \right)^2.$$

4. The facilitator chooses a hypothetical value of Y at $\mathbf{x} = \mathbf{x}_2$. We denote this value by y_2^0. Following the condition specified in step 1, this hypothetical value must not equal $C(Y_2|y_1^0)$.

5. The facilitator elicits the expert's median values of $Y_2, \ldots, Y_r$, conditional on both hypothetical values y_1^0 and y_2^0. The conditional center $C(Y_i|y_1^0, y_2^0)$ required in (8) and (9) is set equal to the elicited conditional median for Y_i.

Design point	Initial percentiles		Hypothetical value (y_i^0)	Conditional assessments						Percentile
	50th	75th		1	2	3	4	5	6	
$\mathbf{x}_1$	385	410	417	390	205	580	265	705	395	50th
					240					75th
$\mathbf{x}_2$	200	235	259		210	585	270	710	400	50th
						615				75th
$\mathbf{x}_3$	575	605	627			590	275	715	405	50th
							310			75th
$\mathbf{x}_4$	260	295	306				280	720	410	50th
								755		75th
$\mathbf{x}_5$	700	745	775					730	415	50th
									455	75th
$\mathbf{x}_6$	390	430	471						425	50th

Table 3: The elicited conditional assessments required to estimate ω and R.

6. The facilitator elicits the expert's 75th percentile for Y_3 conditional on y_1^0 and y_2^0. Denoting this elicited percentile by $y^*_{3,0.75}$, the conditional spread $S(Y_3|y_1^0, y_2^0)$ required in (10) is set equal to

$$\left(\frac{y^*_{3,0.75} - C(Y_3|y_1^0, y_2^0)}{t_{\delta+2}(0.75)}\right)^2.$$

7. This process is repeated for $i = 3, \ldots, r$, where at the i-th step, the facilitator provides hypothetical values $y_1^0, \ldots, y_i^0$, and conditional on these values, the expert gives her medians for $Y_i, \ldots, Y_r$, and her 75th percentile for Y_{i+1}. This continues until U_r has been obtained. (At the r-th step, a 75th percentile is not required.)

Continuing the example, illustrative assessments are given in Table 3.

These are difficult assessments for the expert to make, as it is hard to judge how to change one's beliefs in light of hypothetical data, particularly as this necessarily has to be done without writing down a prior distribution and applying Bayes' theorem. To help both the facilitator and the expert, we suggest the following.

- The facilitator should provide the expert with any graphical displays that might help. We give an example of one such plot in figure 1. The facilitator can report other information that may help, for example, the running mean of the differences $y_i^0 - y_{i,0.5}$, so that the expert can keep track of how much/whether the hypothetical data are consistently above her initial medians.

- The facilitator should discuss with the expert the role of 'noise' in any observed data. Is the expert expecting the error terms to be large (i.e. large σ^2) so that y_i^0 far from $y_{i,0.5}$ is likely? Or is she relatively uncertain about β, such that y_i^0 far from $y_{i,0.5}$ would suggest that her prior assessment of β was wrong and needs revision?
- The facilitator should repeat the exercise for different hypothetical datasets. In the example, the hypothetical data suggested that β_0 had been underestimated. The facilitator could then observe how quickly the expert revised her judgements upwards, as more observations became available. An alternative would have been to give some y_i^0 lower than expected, and some higher, with the implication that some of the effects of the independent variables had been overestimated. The facilitator could then observe to what extent the expert revises her beliefs about the effects of the independent variables, and modifies her assessments accordingly.
- Prior to the elicitation, the facilitator should experiment with different choices of $\mathbf{b}, R, d, \omega$ and hypothetical data (if possible with the expert), so that he has some intuition as to how one 'should' update one's beliefs given new data.

The parameter ω can be estimated from the equations

$$\omega_i = \left\{ S(Y_i|y_1^0, \ldots, y_{i-1}^0) - H_i \right\} \frac{\delta + i - 1}{\delta} F_i,$$

with

$$F_i = \left[1 + \frac{1}{\delta}(y_1^0 - \mathbf{x}_1'\mathbf{b}, \ldots, y_{i-1}^0 - \mathbf{x}_{i-1}'\mathbf{b}) U_{i-1}^{-1} (y_1^0 - \mathbf{x}_1'\mathbf{b}, \ldots, y_{i-1}^0 - \mathbf{x}_{i-1}'\mathbf{b})' \right]^{-1}$$

and

$$H_i = \left[C(Y_i|y_1^0, \ldots, y_i^0) - C(Y_i|y_1^0, \ldots, y_{i-1}^0) \right] \frac{S(Y_i|y_1^0, \ldots, y_{i-1}^0)}{y_i^0 - C(Y_i|y_1^0, \ldots, y_{i-1}^0)}$$

This will give r different estimates $\omega_1, \ldots, \omega_r$ of ω, which can be averaged. Note that ω must be less than or equal to the smallest eigenvalue of U_r, to ensure a valid matrix R. For the example data, we find the smallest eigenvalue of U_r is 1203, with the only one valid estimate of ω 1085 (ω_1 in this case, with $\omega_2, \ldots, \omega_6$ all larger than 1203). As suggested earlier, it is advisable to repeat the elicitation process with a different X in any case, but here the facilitator sets $\omega = 1085$. He now obtains

$$R^{-1} = \begin{pmatrix} 1.5\text{e+}02 & \text{-}3.3\text{e-}04 & \text{-}7.5\text{e-}02 & \text{-}1.7\text{e+}00 & \text{-}1.3\text{e-}01 & \text{-}7.5\text{e+}00 \\ \text{-}3.3\text{e-}04 & 5.0\text{e-}09 & \text{-}1.4\text{e-}07 & \text{-}5.9\text{e-}06 & 3.7\text{e-}07 & 2.6\text{e-}05 \\ \text{-}7.5\text{e-}02 & \text{-}1.4\text{e-}07 & 2.0\text{e-}04 & 2.8\text{e-}03 & 2.9\text{e-}05 & \text{-}4.7\text{e-}03 \\ \text{-}1.7\text{e+}00 & \text{-}5.9\text{e-}06 & 2.8\text{e-}03 & 2.1\text{e-}01 & \text{-}3.8\text{e-}04 & \text{-}2.0\text{e-}01 \\ \text{-}1.3\text{e-}01 & 3.7\text{e-}07 & 2.9\text{e-}05 & \text{-}3.8\text{e-}04 & 1.3\text{e-}04 & 8.8\text{e-}03 \\ \text{-}7.5\text{e+}00 & 2.6\text{e-}05 & \text{-}4.7\text{e-}03 & \text{-}2.0\text{e-}01 & 8.8\text{e-}03 & 1.2\text{e+}00 \end{pmatrix},$$

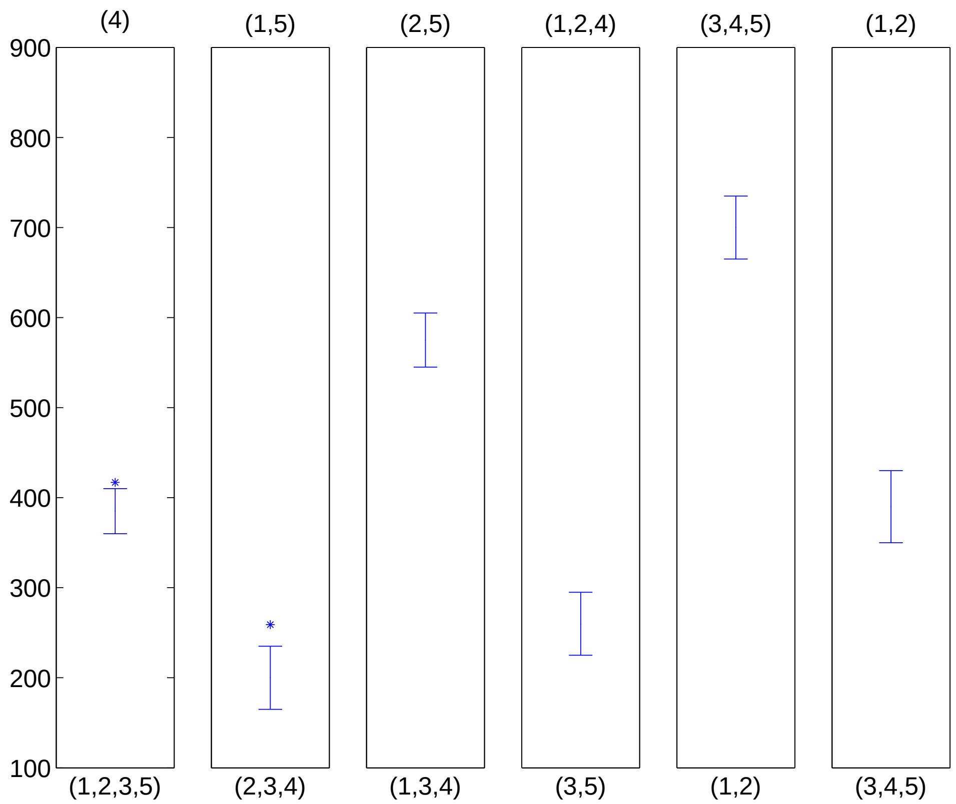

Figure 1: A graph to help the expert make her conditional assessments (corresponding to step 5 in the listed procedure). From left to right, the plots show the expert's original assessments (50% ranges, inferred from her 75th percentiles) for $\mathbf{x}_1$ to $\mathbf{x}_6$ from Table 1. The asterisks show the first two hypothetical observations. The numbers in brackets at the bottom show which inputs are at their 'low' setting, with the numbers at the top showing which inputs are at their 'high' setting. The expert should then re-assess her median for Y_i, for $i = 2, \ldots 6$, and her 75th percentile for Y_3.

completing the (first attempt) at the elicitation of the prior. The facilitator should then report summaries of the distribution of Y for different $\mathbf{x}$ back to the expert, to check that the elicited prior is acceptable. The process can be repeated for a different choice of design matrix X, to check for consistency.

8 Discussion and further reading

The closing discussion from the previous chapter is equally relevant here, and we reiterate the importance of preparation, involving the expert early on in the process, and giving the expert as much training, assistance, and feedback as possible. Not only is multivariate elicitation more time-consuming, but it also typically involves asking harder questions, such as judgements conditional on hypothetical data.

Further discussion of multivariate elicitation is given in Garthwaite et al. (2005) and O'Hagan et al. (2006). One general method, similar in theme to Kadane et al. (1980) is that of probabilistic inversion (see for example Kraan and Bedford, 2005). Again, the idea is to construct a joint distribution for a set of unobservable model parameters or inputs θ from judgements about observable quantities (model outputs) Y. The scenario here is that θ and Y are linked through some deterministic mathematical model $Y = g(\theta)$, where g may be sufficiently complex such that the relationship between θ and Y is not transparent.

Some other topics we have not included in this review are elicitation for logistic regression models (O'Leary et al., 2009), eliciting Dirichlet distributions to represent beliefs about a set of proportions (Chaloner and Duncan, 1987), eliciting beliefs about the parameters of a multivariate normal distribution (Al-Awadhi and Garthwaite, 1998), and eliciting beliefs about correlated binary variables (Papathomas and Hockling, 2003).

References

Al-Awadhi, S. A. and Garthwaite, P. H. (1998). An elicitation method for multivariate normal distributions, *Communications in Statistics - Theory and Methods*, **27**: 1123–1142.

Bedford, T. (2006). Interactive expert assignment of minimally-informative copulae, in *Prcoeedings of the 3rd International Conference on Mathematical Methods in Reliability*, Trondheim, Norway.

Bedford, T. and Cooke, R. M. (2002). Vines - a new graphical model for dependent random variables, *Annals of Statistics*, **30**: 1031–1068.

Bedford, T. and Meeuwissen, A. (1997). Minimally informative distributions with given rank correlation for use in uncertainty analysis, *Journal of Statistical Computation and Simulation*, **57**: 143–174.

Chaloner, K. and Duncan, G. T. (1987). Some properties of the Dirichlet-multinomial distribution and its use in prior elicitation, *Communications in Statistics - Theory and Methods*, **16**: 511–523.

Clemen, R. T., Fischer, G. W. and Winkler, R. L. (2000). Assessing dependence: Some experimental results, *Management Science*, **46**: 1100–1115.

Clemen, R. T. and Reilly, T. (1999). Correlations and copulas for decision and risk analysis, *Management Science*, **45**: 208224.

Daneshkhah, A. and Bedford, T. (2010a). Approximating multivariate distributions with vines, Tech. rep., University of Strathclyde.

Daneshkhah, A. and Bedford, T. (2010b). Eliciting multivariate probability distributions: Copulas and correlations, in *Wiley Encyclopedia of Operations Research and Management Science*, Wiley, under review.

Davies, J., Finlay, M., Mclenaghen, T. and Wilson, D. (2006). Key risk indicators: Their role in operational risk management and measurement, in *The Advanced Measurement Approach to Operational Risk*, edited by E.Davies, Risk Books, London.

De Finetti, B. (1975). *Theory of Probability*, vol. 2, New York: Wiley.

Fackler, P. L. (1991). Modeling interdependence: An approach to simulation and elicitation, *American Journal of Agricultural Economics*, **73**: 1091–1097.

Garthwaite, P. H. and Dickey, J. M. (1988). Quantifying expert opinion in linear regression problems, *J. Roy. Statist. Soc. Ser. B*, **50**: 462–474.

Garthwaite, P. H., Kadane, J. B. and O'Hagan, A. (2005). Statistical methods for eliciting probability distributions, *J. Am. Statist. Assoc.*, **100**: 680–701.

Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S. and Peters, S. C. (1980). Interactive elicitation of opinion for a normal linear model, *J. Am. Statist. Assoc*, **75**: 845–854.

Kadane, J. B. and Wolfson, L. J. (1998). Experiences in elicitation (with discussion), *The Statistician*, **47**: 3–19.

Kraan, B. and Bedford, T. (2005). Probabilistic inversion of expert judgments in the quantification of model uncertainty, *Management Science*, **51**: 995–1006.

Kurowicka, D. and Cooke, R. M. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*, Chichester: Wiley.

Lewandowski, D. (2008). *High Dimensional Dependence: Copulae, Sensitivity, Sampling*, Ph.D. thesis, Delft University.

Moala, F. M. and O'Hagan, A. (2010). Elicitation of multivariate prior distributions: a nonparametric bayesian approach, *J. Statist. Plan. and Infer.*, **140**: 1635–1655.

Morales, O., Kurowicka, D. and Roelen, A. (2008). Eliciting conditional and unconditional rank correlations from conditional probabilities, *Reliability Engineering and System Safety*, **93**: 699–710.

Nixon, R. M., O'Hagan, A., Oakley, J. E., Madan, J., Stevens, J. W., Bansback, N. and Brennan, A. (2009). The rheumatoid arthritis drug development model: A case study in Bayesian clinical trial simulation, *Pharmaceutical Statistics*, **8**: 371–389.

O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications (with discussion), *The Statistician*, **47**: 21–35.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D., Oakley, J. E. and Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*, Chichester: Wiley.

O'Leary, R., Low Choy, S., Murray, J. V., Kynn, M., Denham, R., Martin, T. G. and Mengersen, K. (2009). Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby *Petrogale penicillata*, *Environmetrics*, **20**: 379–398.

Papathomas, M. and Hockling, R. J. (2003). Bayesian updating for binary variables: an application in the UK water industry, *The Statistician*, **52**: 483–499.

Percy, D. F. (2004). Subjective priors for maintenance models, *Journal of Quality in Maintenance Engineering*, **10**: 221–227.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut Statistique de l'Université de Paris*, **8**: 229–231.

Stevenson, M. D., Oakley, J. E., Chick, S. and Chalkidou, K. (2009a). The cost-effectiveness of surgical instrument management policies to reduce the risk of vCJD transmission to humans, *Journal of the Operational Research Society*, **60**: 506–518.

Stevenson, M. D., Oakley, J. E., Lloyd Jones, M., Brennan, A., Compston, J. E., McCloskey, E. and Selby, P. L. (2009b). The cost-effectiveness of an RCT to establish whether 5 or 10 years of bisphosphonate treatment is the better duration for women with a prior fracture, *Medical Decision Making*, **29**: 678–689.