**Revista Colombiana de Estadística - Applied Statistics**
January 2021, volume 44, issue 1, pp. 159 a 170
http://dx.doi.org/10.15446/rce.v44n1.83525

# Elicitation of the Parameters of Multiple Linear Models

## Elicitación de los parámetros de un modelo de regresión lineal múltiple

Juan Carlos Correa-Morales[1,a], Carlos Barrera-Causil[2,b]

[1]Escuela de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, Medellín, Colombia

[2]Grupo de Investigación Davinci, Facultad de Ciencias Exactas y Aplicadas, Instituto Tecnológico Metropolitano, Medellín, Colombia

**Abstract**

Estimating the parameters of a multiple linear model is a common task in all areas of sciences. In order to obtain conjugate distributions, the Bayesian estimation of these parameters is usually carried out using noninformative priors. When informative priors are considered in the Bayesian estimation an important problem arises because techniques are required to extract information from experts and represent it in an informative prior distribution. Elicitation techniques can be used for such purpose even though they are more complex than the traditional methods.

In this paper, we propose a technique to construct an informative prior distribution from expert knowledge using hypothetical samples. Our proposal involves building a mental picture of the population of responses at several specific points of the explanatory variables of a given model and indirectly eliciting the mean and the variance at each of these points. In addition, this proposal consists of two steps: the first step describes the elicitation process and the second step shows a simulation process to estimate the model parameters.

***Key words*:** Bayesian statistics; Conjugate distribution; Elicitation; Informative distribution.

**Resumen**

La estimación de los parámetros de un modelo de regresión lineal múltiple es una tarea común en todas las áreas de las ciencias. Con la idea de obtener distribuciones conjugadas, la estimación Bayesiana de estos parámetros se lleva a cabo usando distribuciones a priori no informativas. Un problema

[a]Ph.D. E-mail: jccorrea@unal.edu.co
[b]Ph.D. E-mail: carlosbarrera@itm.edu.co

importante resulta cuando se incorporan distribuciones a priori informativas en la estimación Bayesiana, puesto que se hace necesario usar técnicas para extraer información de expertos, y representar dicha información en una distribución a prior informativa. Así, los métodos de elicitación pueden ser implementados para tal fin, a pesar de la complejidad de esta tarea en relación con las metodologías tradicionales.

En este paper, se propone un técnica para construir una distribución a priori informativa a partir de muestras hipotéticas usando información de expertos. Esta propuesta se basa en la construcción de un mapa mental de la población de respuestas en diferentes valores específicos de la variable explicativa en el modelo, y luego elicitar de forma indirecta la media y la varianza en cada uno de dichos valores de interés.

La propuesta es presentada en dos pasos, el primer paso describe el proceso de elicitación, y el segundo paso muestra un proceso de simulación para estimar los parámetros del modelo.

***Palabras clave*:** Distribución conjugada; Distribución informativa; Elicitación; Estadística Bayesiana.

# 1. Introduction

One of the main characteristics of the Bayesian inference approach is the incorporation of a prior distribution into the analysis. For this reason, incorporating a prior distribution that actually represents a prior belief about the parameters under study is paramount, so that elicitation can play an important role in obtaining a prior distribution that captures previous beliefs about the parameters and may be used in Bayesian analysis.

Elicitation allows experts to quantify their personal knowledge via probability distributions. This technique has been used in complex problems, where traditional methods (which use noninformative priors) are not applicable or do not perform well (Andrade and Gosling, 2011; Aaron et al., 1995; Truong and Heuvelink, 2013; Nemet et al., 2013; Wilcox et al., 2016). Unfortunately, this complex task has been ignored by many analysts who prefer working with noninformative prior distributions. According to Biedermann et al. (2017), "in essence, a probability expresses a reasoner's uncertainty about something—for example a state of nature of the past, present or future—that is not completely known to this person." For their part, O'Hagan and Oakley (2004) and Kadane and Wolfson (1998) discuss the difficulties of the elicitation process.

Proportion is the parameter that has received the most attention in the elicitation process in both nonparametric and parametric approaches (Winkler, 1967*a*,*b*; Raiffa, 1970; Chaloner and T., 1983; Gavasakar, 1988; Umesh, 1988). In his study, Tversky (1974) addresses the problems that arise when eliciting subjective probabilities. Gzyl et al. (2017) consider the problem of experts providing intervals of uncertainty rather than point estimates of the parameters of interest. Shadbolt and Burton (1995) describe in detail the steps to execute an elicitation process. O'Hagan (2019) presents a complete description of a number of best practices through which elicitation can be made as rigorous and scientific as possible in order to minimize the cognitive biases that experts tend to have when making probabilistic judgments.

Several techniques to help experts elicit their personal knowledge in terms of probability distributions have been proposed, such as the betting- and lottery-based methods (Holloway, 1979; Harrison et al., 2014). These two methods require a very strong assumption of linearity of the personal utility function (DeGroot, 1970). Other methods resort to a more intuitive approach, such as creating a mental picture of the population. This allows experts to obtain mental samples of a population that will hopefully contain a sample that preserves its structure. The quality of its personal information is quantified in terms of the equivalent sample size. Christov et al. (2017) apply several elicitation techniques; however, their study does not provide general conclusions given its extreme limitations. Renooij and Witteman (1999) and Witteman and Renooij (2003) propose using verbal expression to quantify personal probabilities. Furthermore, although different software tools for Expert Knowledge Elicitation (EKE) have been developed, none of them has executed, or simplified, parameter elicitation in linear models (Fisher et al., 2012; Seynaeve et al., 2019).

One of the main difficulties when eliciting expert knowledge is to select and formulate the appropriate questions in such a way that experts understand what is being asked and actually provide the expected information (Andrade and Gosling, 2018). For this purpose, a list of questions should be carefully prepared. However, during the elicitation process, the structure of a question could be slightly changed, provided that its underlying meaning is not altered. According to Barrera-Causil et al. (2019), numerical skills do not play an important part in the accuracy of the estimation when elicitation is used. The way questions are asked and language is used is considered a major issue because experts may have a low or high level of education. Therefore, all elicitation methods should be examined in order to identify the appropriate vocabulary for the questions and compare all those characteristics based on the different parameters to estimate. In this paper, we use generic and simple questions that experts usually interpret correctly. Nevertheless, in some cases, we must change the words we often use to achieve our final goal. In other words, rather than using structural questions in an elicitation process, we must adapt them to each expert and try to interact with them during the interview/survey process.

An example of a situation in which elicitation would be difficult to apply is when trying to estimate areas with a particular characteristic, such as those that need ecological restoration, based on maps and the knowledge of farmers. In this case, farmers are the experts, but they often have no mathematical or statistical skills. Thus, for a successful elicitation, this process must be adapted to their abilities in such a way that they can provide relevant information.

In this study, the elicitation process is based on the construction of a mental picture (or mental model) of the population of responses at several specific points of the explanatory variables in a linear model. Hence, we define a mental model as a representation of the way human beings see the world based on their lessons learned and their interaction with the things around them. In his studies, Johnson-Laird (2010), Johnson-Laird (1994), and Johnson-Laird (1980) provides a comprehensible description of what a mental model is. He describes, in a deeper sense, how human reasoning, probabilistic thinking, and cognitive processes act to reproduce a mental model.

A simple example of a mental model could be a witness' image of the appearance or physical features of a criminal. That graphical representation would be the mental model of the witness, which is constructed based on his/her experience. Thus, expert knowledge elicitation via the mental model involves extracting all the experts' information about a parameter of interest. This process considers how experts are able to express their probabilistic thinking and how their reasoning and cognitive processes could influence a good estimation of such parameter.

Eliciting an unknown vector parameter of a normal linear model is really important. In particular, James et al. (2010) developed a software tool to help with this task. However, one of the drawbacks of their study and software is that it requires a direct elicitation of the model parameters. This could be easier if there were previous studies providing direct information about them.

Our proposal involves constructing a mental picture of the population of responses at several specific points of the explanatory variables in a linear model and indirectly eliciting the mean and the variance at each of these points. This task is by far easier than the direct elicitation of the model parameters of a linear model and represents the first step of the algorithm shown in Section 2. Also, Section 2 presents the second step of the algorithm. This step includes a simulation to generate samples of the response variable and repeatedly estimate the model parameters in order to obtain the conjugate prior distribution of a regression model. Section 3 provides a hypothetical example, in which the parameters of a multiple linear model are estimated. Finally, section 4 draws the conclusion of the study.

## 2. The Algorithm

Let us assume, by simplification, that we are interested in the simple linear model presented below. Note that this proposal may be extended to multiple linear regression, as reported in the hypothetical example in Section 3. However, if the number of model parameters increase, the elicitation process could become exhausting. Then, let

$$y \sim \beta_0 + \beta_1 \cdot x + \epsilon,$$

where $\epsilon$ is assumed to follow a normal distribution with mean 0 and precision $\tau = \frac{1}{\sigma^2}$. Some Bayesian statisticians prefer to use the inverse of the variance, which they call precision (DeGroot, 1970). This facilitates the algebraic work to get conjugate distributions, working with the inverse gamma distribution instead of the traditional gamma distribution. Our basic problem here is to determine the conjugate prior joint distribution of the model parameters, say $\xi\left(\beta_0, \beta_1, \sigma\right)$.

Now, we present our two-step proposal. In the first step, some quantiles of the distribution of the response variable in the regression model are elicited with respect to a particular value of the explanatory variable. The mean, the variance, and the corresponding equivalent sample size are also obtained in this step. Note that eliciting the quantiles of a random variable is easier that eliciting the model parameters directly (Demuynck, 2013).

In the second step, a simulation is performed to generate samples of the response variable, considering their corresponding design points, in order to repeatedly estimate the model parameters several times. It is easier for experts to provide their opinions around a central value than around any other value in the range of possible values of a variable. This argument is based on the fact that people find it easier to describe events that occur more frequently than those that are rare. For this reason, it is recommended to establish, as design points, values close to the mean of the explanatory variable. At this point, any type of bias (e.g., anchoring, availability, or overconfidence) during elicitation should be avoided, as it may lead the elicitation process on the wrong track and any implemented method would fail to make a good estimate of the prior distributions (O'Hagan, 2019).

***Step 1:*** *Elicitation*

First, the expert is asked to choose a value for the explanatory variable within its logical range, say $x_0$. Then, he/she is asked to imagine the distribution of the responses at this level of the explanatory variable, $x_0$, and to select a value for the response within its logical range, say $y_0$.

Subsequently, the expert is presented with the following situation: If we had a representative sample of a conditional population of size $n_0$, how many observations would you believe are below $y_0$? Let us call this number $n_{y0}$. Then, the expert is asked for the minimum number of observations in the sample that he/she would accept below $y_0$ (let us call this number $k_l$) and for the maximum number of observations in the sample that he/she would accept above $y_0$ (lest us call this number $k_u$).

At each elicited value, the expert must produce a hypothetical sample size that reflects his/her self-confidence in this elicited distribution. At the end of the process, we will have

$$\left\{\xi\left(y_i \left| x_{i1}, x_{i2}, \cdots, x_{ip}\right.\right), n_i\right\}_{i=1.}^{k}$$

Let us suppose that we obtain the following values:

Question 1:

| Level of the explanatory variable ($X = x_0$) | | | |
|---|---|---|---|
| | Hypothetical sample | Number given | Expected percentile |
| The most likely | $n_0$ | $n_{y_0}$ | $E\left[Percentile\right] = \frac{n_{y_0}}{n_0}$ |
| More likely | $n_0$ | $n_{y_0} + k_u$ | $E_u\left[Percentile\right] = \frac{n_{y_0}+k_u}{n_0}$ |
| The least likely | $n_0$ | $n_{y_0} - k_l$ | $E_l\left[Percentile\right] = \frac{n_{y_0}-k_l}{n_0}$ |

Question 2:

| Level of the explanatory variable ($X = x_0$) | | | |
|---|---|---|---|
| | Hypothetical sample | Number given | Expected percentile |
| The most likely | $n_1$ | $n_{y_1}$ | $E\left[Percentile\right] = \frac{n_{y_1}}{n_1}$ |
| More likely | $n_1$ | $n_{y_1} + k_u$ | $E_u\left[Percentile\right] = \frac{n_{y_1}+k_u}{n_1}$ |
| The least likely | $n_1$ | $n_{y_1} - k_l$ | $E_l\left[Percentile\right] = \frac{n_{y_1}-k_l}{n_1}$ |

Now, let us assume that, for the same value of explanatory variable $x_0$ and under $n_0$, we can standardize $y_0$ such that

$$\frac{y_0 - \mu_{x_0}}{\sigma} \approx \Phi^{-1}\left(\frac{n_{y_0}}{n_0}\right).$$

Similarly, but under $n_1$, we have

$$\frac{y_1 - \mu_{x_0}}{\sigma} \approx \Phi^{-1}\left(\frac{n_{y_1}}{n_1}\right).$$

Then, we solve the following equations for $\mu_{x_0}$ and $\sigma$:

$$\sigma = \frac{y_0 - y_1}{\Phi^{-1}\left(\frac{n_{y_0}}{n_0}\right) - \Phi^{-1}\left(\frac{n_{y_1}}{n_1}\right)}$$

$$\mu_{x_0} = y_0 - \left(\frac{y_0 - y_1}{\Phi^{-1}\left(\frac{n_{y_0}}{n_0}\right) - \Phi^{-1}\left(\frac{n_{y_1}}{n_1}\right)}\right)\Phi^{-1}\left(\frac{n_{y_0}}{n_0}\right).$$

Next, we must find an equivalent sample size ($n_{equ}$) that reflects the expert's self-confidence in the information they provided. For this purpose, we could use the following binomial confidence interval for a proportion:

$$\hat{\pi} \pm 1.96\sqrt{\frac{\hat{\pi}\left(1 - \hat{\pi}\right)}{n_{equ}}}.$$

If we assume that the upper limit is $\frac{n_{y_0}+k_u}{n_0}$, then

$$\frac{n_{y_0} + k_u}{n_0} = \frac{n_{y_0}}{n_0} + 1.96\sqrt{\frac{\frac{n_{y_0}}{n_0}\left(1 - \frac{n_{y_0}}{n_0}\right)}{n_{equ}}}.$$

Thus, the equivalent ($n_{equ}$) to the expert's self-confidence at this design point is approximated as follows:

$$n_{equ1} = \left(\frac{1.96}{k_u}\right)^2 n_{y_0}\left(n_0 - n_{y_0}\right).$$

If we assume that the lower limit is $\frac{n_{y_0}-k_l}{n_0}$, then

$$\frac{n_{y_0}-k_l}{n_0} = \frac{n_{y_0}}{n_0} - 1.96\sqrt{\frac{\frac{n_{y_0}}{n_0}\left(1-\frac{n_{y_0}}{n_0}\right)}{n_{equ}}}.$$

Thus, the equivalent ($n_{equ}$) to the expert's confidence at this design point is approximated as follows:

$$n_{equ2} = \left(\frac{1.96}{k_l}\right)^2 n_{y_0}\left(n_0 - n_{y_0}\right).$$

Since we would rather be cautious regarding the expert's information, we choose the following equivalent sample size: $n_{x_0} = \min\{n_{equ1}, n_{equ2}\}$.

A similar process is considered, assuming $x_1$ is a different level of the explanatory variable.

In summary, note that, for two points of the explanatory variable, say $x_0$ and $x_1$, we have two tables as the ones presented above. From the two "Most likely" rows, we obtain the following two equations:

$$\begin{aligned} \frac{x_0-\mu}{\sigma} &= \Phi^{-1}\left(\frac{n_{x_0}}{n_0}\right), \text{ and} \\ \frac{x_1-\mu}{\sigma} &= \Phi^{-1}\left(\frac{n_{x_1}}{n_1}\right). \end{aligned}$$

To clarify the elicitation process described above (in which we need to obtain specific information from experts), let us suppose that we have a population of soccer players, and our purpose is to model their weight ($y_0$) using their height ($x_0$) as the explanatory variable. Then, if we have a sample of size $n_0$ which includes soccer players with $x_0 = 170$ cm, we could ask the expert the following: how many soccer players, at most, would you believe weigh less than 70 kg? His/Her answer would be $n_{y_0}$. If we ask the expert what he/she would consider to be the maximum acceptable number of soccer players in that particular sample that weigh 70 kg, his/her answer would be $k_u$. Now, if we ask the expert what he/she would consider to be the minimum possible number of soccer players in that sample that weigh less than 70 kg, his/her answer would be $k_l$.

Likewise, if we use another hypothetical sample, say $n_1$, the questions could be as follows: in a sample of soccer players who are 170 cm tall, how many of them, at most, would you believe weigh less than 80 kg? What would you consider to be the maximum acceptable number of players in that particular sample that weigh 80 kg? What would you consider to be the minimum possible number of players in that sample that weigh less than 80 kg? This process can be repeated as many times as the researcher deems necessary.

Once we have the populations at the several proposed points of the explanatory variables, along with their respective equivalent sample sizes, we proceed to construct the conjugate prior distribution of the regression parameters.

***Step 2:*** *Simulation*

1. For the $i$-th population specified above, obtain a simulated sample of size $n_{x_i}$, say $\mathbf{y}_i$.
2. Construct

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$$

and

$$\mathbf{X} = \begin{bmatrix} \mathbf{1} & \mathbf{x}_1 \\ \mathbf{1} & \mathbf{x}_2 \end{bmatrix}.$$

3. Estimate the model $y = \beta_0 + \beta_1 x + e$, where $E\left[e\right] = 0$ and $Var\left(e\right) = \sigma^2$. Keep $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2\right)$.
4. Repeat the above steps several times, say $M$.

After the above process is completed, we estimate the parameters of a multinormal distribution for the $\beta$'s and a Gamma for the variance. The conjugate prior is the normal-inverse-gamma distribution, which is expressed as follows:

$$\begin{aligned} \xi\left(\beta, \sigma^2\right) &= \xi(\beta \mid \sigma^2)\xi(\sigma^2) = N(\mu_\beta, \sigma^2\Sigma_\beta) \times IG(\alpha_0, \gamma_0) \\ &\propto \left(\frac{1}{\sigma^2}\right)^{(\alpha_0 + p/2 + 1)} \times \exp\left[-\frac{1}{\sigma^2}\left\{\gamma_0 + \frac{1}{2}(\beta - \mu_\beta)'\Sigma_\beta(\beta - \mu_\beta)\right\}\right], \end{aligned}$$

where $p$ is the number of regressors, and the hyperparameters $\alpha_0$ and $\gamma_0$ should be estimated using a noninformative prior or through elicitation.

Determining $\alpha_0$ and $\beta_0$ is often surrounded by uncertainty. In this case, researchers place a distribution on $(\alpha_0, \gamma_0)'$, which is known as a hyper-prior distribution.

Finally, with the elicitation method mentioned above, Bayesian inference can be performed, considering informative priors.

# 3. Hypothetical Example

To illustrate this procedure, let us assume that we are interested in modeling the weight of undergraduate students based on their height and sex using a linear model without interaction.

$$Weight = \beta_0 + \beta_1 Sex + \beta_2 Height + e,$$

where $e \sim N\left(0, \sigma^2\right)$. If we choose four points to elicit the distribution of the students' weight (two points for each sex, say $x_{1M} = 170$ cm and $x_{2M} = 180$ cm, and $x_{1F} = 160$ cm and $x_{2F} = 170$ cm), we elicit the parameters of four population normal distributions; here,eliciting the mean and the variance of each distribution can be easier than eliciting the original parameters of the model.

We could ask questions to the expert using each design point and different hypothetical samples, as follows:

**For male students:** In a sample of size $n_0$ which includes undergraduate students that are 170 cm tall, how many students at most would you believe weigh less than 70 kg? What would you consider to be the maximum acceptable number of students in that particular sample that weigh 70 kg? What would you consider to be the minimum possible number of students in that sample that weigh less than 70 kg?

Now, using the two tables, as in the previous examples, from the two "Most likely" rows, we could obtain the mean and variance of each population and generate the corresponding $y_{ki}$, $k = 1, 2, 3, 4$, and $i = 1, 2, \ldots, m$ from $N_{\{M,x=170\}}\left(\mu_1\,,\,\sigma_1^2\right)$, $N_{\{M,x=180\}}\left(\mu_2\,,\,\sigma_2^2\right)$, $N_{\{F,x=160\}}\left(\mu_3\,,\,\sigma_3^2\right)$, and $N_{\{F,x=170\}}\left(\mu_4\,,\,\sigma_4^2\right)$.

Now, with

$$\mathbf{Y} = \begin{bmatrix} y_{11} \\ \vdots \\ y_{n_{x_0}1} \\ y_{12} \\ \vdots \\ y_{n_{x_1}2} \\ y_{13} \\ \vdots \\ y_{n_{x_2}3} \\ y_{14} \\ \vdots \\ y_{n_{x_3}4} \end{bmatrix}, \mathbf{X} \begin{bmatrix} 1 & 1 & 170 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 170 \\ 1 & 1 & 180 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 180 \\ 1 & 0 & 160 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 160 \\ 1 & 0 & 170 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 170 \end{bmatrix}.$$

Next, we estimate $(\beta_0, \beta_1, \beta_2, \sigma^2)$ following step 3 of the algorithm. Finally, with steps 2, 3, and 4, we obtain the conjugate prior distribution.

# 4. Conclusions

Since elicitation is an extremely complex and undervalued process, further research is needed to scientifically compare the different elicitation methods that have been proposed. Psychologists and cognitive scientists are required to thoroughly study the process that takes place when an expert is faced with a particular elicitation task.

In this study, we proposed using an algorithm based on a hypothetical mental sample to determine the parameters of the conjugate distribution of a normal population with unknown parameters. In our experience, experts feel more comfortable when they are asked questions about quantiles of this type of hypothetical sample. Moreover, it is easy to extend this algorithm to other situations, including more complex models such as regression ones. The R code of the algorithm for a linear model with two design points is provided at `https://github.com/cbarrera2101/Code-elicitation-of-the-parameters-for-LM.git`

Our proposal could be easily applied to multiple linear models with categorical variables, such as independent and continuous variables, or a mix of both. However, the response variable should always be quantitative.

# Acknowledgements

C.B.-C. thanks Instituto Tecnológico Metropolitano -ITM- for their support.

[Received: March 2019 — Accepted: December 2020]

# References

Aaron, R., DeWispelare, A. R., Herren, L. T. & Clemen, R. T. (1995), 'The use of probability elicitation in the high-level nuclear waste regulation program', *International Journal of Forecasting* **11**, 5–24.

Andrade, J. A. A. & Gosling, J. P. (2011), 'Predicting rainy seasons: quantifying the beliefs of prophets', *Journal of Applied Statistics* **38**(1), 183–193.

Andrade, J. A. A. & Gosling, J. P. (2018), 'Expert knowledge elicitation using item response theory', *Journal of Applied Statistics* **45**(16), 2981–2998.

Barrera-Causil, C. J., Correa, J. C. & Marmolejo-Ramos, F. (2019), 'Experimental investigation on the elicitation of subjective distributions', *Frontiers in Psychology* **10**, 862.

Biedermann, A., Bozza, S., Taroni, F. & Aitken, C. (2017), 'The consequences of understanding expert probability reporting as a decision', *Science and Justice* **57**, 80–85.

Chaloner, K. & T., D. (1983), 'Assessment of a beta prior distribution: Pm elicitation', *The Statistician* **27**, 174–180.

Christov, S. C., Marquard, J. L., S., G., Avrunin, G. S. & Clarke, L. A. (2017), 'Assessing the effectiveness of five process elicitation methods: A case study of chemotherapy treatment plan review', *Applied Ergonomics* **59**, 364–376.

DeGroot, M. H. (1970), *Optimal Statistical Decisions*, McGraw Hill, New York.

Demuynck, T. (2013), 'A mechanism for eliciting the mean and quantiles of a random variable', *Economics Letters* **121**(1), 121–123.

Fisher, R., O'Leary, R. A., Low-Choy, S., Mengersen, K. & Caley, M. J. (2012), 'A software tool for elicitation of expert knowledge about species richness or similar counts', *Environmental Modelling and Software* **30**, 1–14.

Gavasakar, U. (1988), 'A comparison of two elicitation methods for a prior distribution for a binomial parameter', *Managment Science* **34**(6), 784–790.

Gzyl, H., ter Horst, E. & Molina, G. (2017), 'Inferring probability densities from expert opinion', *Applied Mathematical Modelling* **43**, 306–320.

Harrison, G. W., Martínez-Correa, J. & Swarthout, J. T. (2014), 'Eliciting subjective probabilities with binary lotteries', *Journal of Economic Behavior & Organization* **101**, 128–140.

Holloway, C. A. (1979), *Decison Making Under Uncertainty: Models and Choices*, Prentince-Hall, Inc., Englewood Cliffs, NJ.

James, A., Choy, S. L. & Mengersen, K. (2010), 'Elicitator: An expert elicitation tool for regression in ecology', *Environmental Modelling & Software* **25**, 129–145.

Johnson-Laird, P. (1980), 'Mental models in cognitive science', *Cognitive Science* **4**(1), 71–115.

Johnson-Laird, P. N. (1994), 'Mental models and probabilistic thinking', *Cognition* **50**, 189–209.

Johnson-Laird, P. N. (2010), 'Mental models and human reasoning', *Proceedings of the National Academy of Sciences* **107**(43), 18243–18250.

Kadane, J. B. & Wolfson, L. J. (1998), 'Experiences in elicitation', *The Statistician* **47**(1), 3–19.

Nemet, G. F., Baker, E. & Jenni, K. E. (2013), 'Modeling the future costs of carbon capture using experts? elicited probabilities under policy scenarios', *Energy* **53**, 218–228.

O'Hagan, A. (2019), 'Expert knowledge elicitation: Subjective but scientific', *The American Statistician* **73**(sup1), 69–81.

O'Hagan, A. & Oakley, J. E. (2004), 'Probability is perfect, but we can't elicit it perfectly', *Reliability Engineering and System Safety* **85**, 239–248.

Raiffa, H. (1970), *Decision Analysis: Introductory Lectures on Choice Under Uncertainty*, Addison-Wesley: Reading, Masschusetts.

Renooij, S. & Witteman, C. (1999), 'Talking probabilities: Communicating probabilistic information with words and numbers', *International Journal of Approximate Reasoning* **22**, 169–194.

Seynaeve, D., Varewyck, M. & Verbeke, T. (2019), 'Extension of the monte-carlo web application and expert knowledge elicitation web application', *EFSA Supporting Publications* **16**(6), 1630E.

Shadbolt, N. & Burton, M. (1995), 'Knowledge elicitation: A systematic approach', *Evaluation of Human Work: A Practical Ergonomics Methodology* pp. 406–440.

Truong, P. N. & Heuvelink, G. B. M. (2013), 'Uncertainty quantification of soil property maps with statistical expert elicitation', *Geoderma* **202-203**, 142–152.

Tversky, A. (1974), 'Assessing uncertainty', *Journal of the Royal Statistical Society. Series B (Methodological)* **36**(2), 148–159.

Umesh, G., A. (1988), 'Comparison of two elicitation methods for a prior for a binomial parameter', *Management Science* **34**, 784–790.

Wilcox, C., Mallos, N. J., Leonard, G. H., Rodriguez, A. & Hardesty, B. D. (2016), 'Using expert elicitation to estimate the impacts of plastic pollution on marine wildlife', *Marine Policy* **65**, 107–114.

Winkler, R. L. (1967*a*), 'The assessment of prior distributions in bayesian analysis', *Journal of the American Statistical Association* **62**(319), 776–800.

Winkler, R. L. (1967*b*), 'The quantification of judgement: Some methodological suggestions', *Journal of the American Statistical Association* **62**(320), 1105–1120.

Witteman, C. & Renooij, S. (2003), 'Evaluation of a verbal numerical probability scale', *International Journal of Approximate Reasoning* **33**, 117–131.