A Protocol for the Elicitation of Prior Distributions

Gertrudes Coelho Nadler Lins
Mestrado em Engenharia de Producao
Universidade Federal de Pernambuco
Recife / PE, Brazil
gertie@uol.com.br

Fernando Menezes Campello de Souza
Departamento de Eletronica e Sistemas
Universidade Federal de Pernambuco
Recife / PE, Brazil
fmcs@elogica.com.br

Abstract

A practical way of eliciting convex sets of probability measures on a real continuous variable is introduced, one which mathematically defines vagueness and allows for its explicit treatment when it emerges from the activity of making inferences about a parameter based on available evidence through expert opinion. In the setup of the protocol, new indexes are introduced concerning the detailing of the questionnaire.

Keywords. elicitation, uncertainty, prior knowledge, prior distribution, expert opinion, convex sets of probability measures, vagueness.

1 Introduction

Bayesian inference provides a method for the treatment of the *a priori*, subjective, accumulated knowledge that one has about a state of the world. The *a priori* probability, used in Bayesian inference, is also called subjective or epistemic probability, and it represents the degree of belief that the individual has in the occurrence of an event which is represented in terms of the variable θ. Notwithstanding its claimed advantages, one of the main drawbacks of this approach is that a precise prior distribution is needed. The attempts to circumvent this disadvantage include the consideration of families of prior distributions, as in sensitivity analysis or robustness. More generally, there is a whole field of *imprecise probabilities* that deals with this kind of problem.

The protocol presented here is part of a method that provides a systematic procedure for the elicitation of a *a priori* knowledge, i.e., of a prior distribution of some unknown real-valued continuous parameter, from an expert. It is useful in practical, real world, settings, where, typically, the available evidence is of a mixed, partial, nature, and the expert's knowledge has always a certain degree of vagueness.

The protocol is based on the research results presented in [1], [2] and [3]. It avoids betting schemes, which are a source of confusion, given that judgments would be elicited through preferences, and this would involve two different psychological mechanisms. Also, it does not require total precision on behalf of the specialist, and there are no errors to be treated statistically. The method uses paired comparative probabilistic assertions involving intervals of θ values, and the choice of the events (intervals of θ values) for the paired comparisons and the paired comparisons themselves (the items in the questionnaire), are made according to new constructs that are introduced here. More material on these sorts of ideas may be found in references [4] to [13].

2 The Method

The general method is introduced in [13], but let us briefly describe it here. For the elicitation procedure, the first assumption in the development of the method is that the specialist has an incomplete knowledge about the probability distribution on θ, i.e., about the prior distribution $\pi(\theta)$. More specifically, it is assumed that the specialist can make only finitely many comparative probabilistic assertions in answer to questions about the likelihood of θ belonging to one of two given intervals. The basis for such assumption are the feasibility of a reasonable protocol (questionnaire), and the natural limitations of human beings. Thus, the specialist can have his or her prior knowledge represented by a family of probability distributions containing a distribution that is stochastically "larger" than all the other distributions compatible with the answers that were given, and also by a distribution that is stochastically "smaller" than all the others. This family of distributions would be formed by the set of all convex combinations of the two cited distributions. The extreme distributions are equivalent to a minimum (maximum) expected value probability distribution.

Initially, an elicitation questionnaire has to be set up. Consider the case of a continuous real parameter. First, minimum and maximum plausible values for θ (that is, $\theta_{\min}$ and $\theta_{\max}$) have to be established in such a way that the probability that the true value of θ lies outside these two limits is zero. Those bounds should be specified by the specialist. It is necessary that, without loss of generality, the probability that θ belongs to any subinterval of the interval $[\theta_{\min},\theta_{\max}]$ should be different from zero. Then, the probability that θ belongs to this interval is equal to one. Inside this interval, θ is distributed according to an unknown probability density $\pi(\theta)$. The interval $[\theta_{\min},\theta_{\max}]$ is partitioned into 2n subintervals of equal Lebesgue measure, $[\theta_{j-1},\theta_j), j = 1,2,...,2n-1$; $[\theta_{2n-1},\theta_{2n}]$. Define also $\pi_j = \Pr\{\theta \in [\theta_{j-1},\theta_j)\}$, the probability that θ belongs to the jth subinterval. The probability that θ belongs to the interval $[\theta_j,\theta_{j+k})$ is $\sum_{i=1}^{k}\pi_{j+i}$. It is clear that $\sum_{j=1}^{2n}\pi_j = 1$.

The questions posed to the specialist are of the following type: Which one is greater than the other,

$\Pr\{\theta \in [\theta_{i-1},\theta_i) \bigcup [\theta_i,\theta_{i+1}) \bigcup ... \bigcup [\theta_{i+k},\theta_{i+k+1})\}$ or

$\Pr\{\theta \in [\theta_{s-1},\theta_s) \bigcup [\theta_s,\theta_{s+1}) \bigcup ... \bigcup [\theta_{s+m},\theta_{s+m+1})\}$?

There will be, then, three possibilities for the answer to each question: greater than, less than, or blank (if the expert cannot compare the two intervals). An answer "equal to" is not considered since this would require from the expert an infinite precision.

The superposition of intervals may cause confusion, so one should formulate questions so that $i+k+1 \leq s-1$, and $i < s$. The first few questions should be the easiest ones, i.e., those involving events that are easily separable as far as their probabilities are concerned (they would involve larger intervals). The first question will be then: Which one is greater,

$\Pr\{\theta \in [\theta_0,\theta_1) \bigcup [\theta_1,\theta_2) \bigcup ... \bigcup [\theta_{n-1},\theta_n)\}$ or

$\Pr\{\theta \in [\theta_n,\theta_{n+1}) \bigcup [\theta_{n+1},\theta_{n+2}) \bigcup ... \bigcup [\theta_{2n-1},\theta_{2n}]\}$?

The intervals are, then, progressively refined. The questions should not be repeated in order to avoid mnemonic "anchoring" effects.

Let I_A and I_B be two given intervals. One could also ask the expert if, besides stating that, say, $P(\theta \in I_A) - P(\theta \in I_B) \leq 0$, he or she could also specify positive numbers a_A and a_B such that $a_A P(\theta \in I_A) - a_B P(\theta \in I_B) \leq 0$.

Having done the above, the following linear programming (LP) problems are posed:

$$Max_{\pi_j}(Min) \sum_{j=1}^{2n} c_j \pi_j \tag{1}$$

subject to:

$$a_{ik}\sum_{j=i}^{k}\pi_j \; - \; a_{lm}\sum_{j=l}^{m}\pi_j \; \leq 0 \tag{2}$$

$(or \geq 0, \; depending \; on \; the \; specialist's \; answer)$

where $k < l$, $a_{ik} > 0$, $a_{lm} > 0$;

$$\alpha_j \pi_j \leq \pi_{j+1}, \; j = 1,2,...,2n-1, \; \alpha_j > 0 \tag{3}$$

$$\beta_j \pi_{j+1} \leq \pi_j, \; j = 1,2,...,2n-1, \; \beta_j > 0 \tag{4}$$

$$\pi_j \geq 0, j = 1,2,...,2n \tag{5}$$

$$\sum_{j=1}^{2n}\pi_j = 1 \tag{6}$$

The input from the specialist consists, then, in answering a certain number of pairwise comparisons of the probabilities of events, and also to express the relative odds.

In restrictions 3 and 4, α_j and β_j could be chosen in such a way as to maximize the entropy of the distribution π on θ, subject to the restrictions representing the specialist's answers. That is, outside the partial prior information available, it is desired sometimes, as exposed in [4], to use a prior that is as uninformative as possible. If that is not desired, one needs only to suppress restrictions 3 and 4.

If it is desirable to get all the probability distributions consistent with the elicited responses, it would be necessary to solve many LP problems where each c_j would be randomly chosen according to a uniform distribution. Or to use any technique to obtain the set of all feasible $\pi_j's$.

If it is desirable to get the distribution with the smallest average value, and the one with the greatest average value, but consistent with the specialist's answers, then one should set $c_j = 2n - j + 1$.

In order to concentrate on the main point, concerning the choice of the intervals for paired comparisons, and the assemblage of the questionnaire, we will consider $c_j = 2n - j + 1$, $a_{ik} = a_{lm} = 1$ in all paired comparisons, and suppress restrictions 3 and 4.

If the specialist could not make a certain paired comparison, the corresponding restriction in expression 2 would be suppressed. If the expert can answer only a few questions, leaving blank most of them, he or she will be in a state of near-ignorance.

If the feasible set of the LP problems turns out to be empty, this means that the specialist is inconsistent.

The common feasible set of these LP problems is defined by the specialist answers and the conditions that guarantee $\{\pi_j\}_{j=1}^{2n}$ to be a probability distribution.

Clearly there is a limit for the perception (precision) of the specialist, so the lengths of the intervals should not be too small. In fact, the specialist should state when his or her limit of precision is reached and no longer can comparisons be made. If the feasible set turns out to be empty, that is, if his or her answers are inconsistent, the specialist should redo the questionnaire and stop after having answered questions of a coarser level immediately before the one he or she initially answered. Or the refined questions should be progressively suppressed until the feasibility of the LP problems is reached. The process is interactive. In any case at a certain coarseness he or she stops.

The specialist's answers will form the so called technological matrix of the two LP problems. If the individual's knowledge is very good, that is to say, if he or she is very precise, the two LP problems will tend to have the same answer, but in general the answers will be different, since there is always vagueness in an individual's knowledge. One will get then two distributions, and any convex combination of them will be feasible. The family of distributions thus obtained will be consistent with the answers of the specialist.

Typically, we will have two probability distributions and the new constructs will be defined in terms of these distribution functions.

2.1 Notation

Parameter (State of Nature): it is represented by θ; it is the "expert inner random variable" (*de dicto*) whose probability distribution will be elicited from the expert himself or herself. It is assumed that θ belongs to a finite closed interval of real numbers.

Parameter Space: $\Theta = \{\theta\} = [\theta_{min}, \theta_{max}]$.

Elementary Interval: It is represented by $\epsilon_j = [\theta_{j-1}, \theta_j)$, $j = 1, 2, ...2n$, where the Lebesgue measure of each one of them will be given by $(\theta_{max} - \theta_{min})/2n$. (Rigorously the last elementary interval would be $[\theta_{2n-1}, \theta_{2n}] = [\theta_{2n-1}, \theta_{max}]$, but since θ is a continuous variable, $Pr\{\theta = \theta_{max} = \theta_{2n}\} = 0$, and so the convenience of the notation will be maintained).

Note that $\theta_{min} = \theta_o$ e $\theta_{max} = \theta_{2n}$.

Partition of the Parameter Space: It is represented by $\wp = \{\epsilon_j\}$, $j = 1, 2, 3, .., 2n$.

Interval: It is represented by $I(k, m) = \bigcup_{j=k}^{m} \epsilon_j = \bigcup_{j=k}^{m} [\theta_{j-1}, \theta_j), k \leq m$, that is, it is the union of contiguous elementary intervals.

Pairs of Intervals (also denominated Questions): They are represented by $P(k, m, s, u) = \{I_A(k, m), I_B(s, u)\} = \{\bigcup_{j=k}^{m} [\theta_{j-1}, \theta_j), \bigcup_{j=s}^{u} [\theta_{j-1}, \theta_j)\}$, for $1 \leq k \leq m < s \leq u \leq 2n$. The set of all such questions is represented by $\Psi = \{P(k, m, s, u)\}$. For simplicity of notation the question involving the smaller values of θ (the left side of the questionnaire) will be represented by I_A, and the one involving the larger values of θ will be denoted by I_B (the right side of the questionnaire). Note that $I_A \cap I_B = \emptyset$.

Questionnaire: It is represented by $Q = \{P_i(k, m, s, u)\}$, for $i = 1, 2, 3, ..., q$, where $P_i(k, m, s, u)$ is a question, as defined above. It is composed, of course, of distinct questions. Note that $Q \subset \Psi$.

3 Vagueness and Precision

The distribution functions on θ, $\Pi_{\max}$ and $\Pi_{\min}$, are easily constructed from the solutions of the respective LP problems. For any $\{\pi_j\}_{j=1}^{2n}$, $\Pi_j = \sum_{i=1}^{j} \pi_i$. If one thinks about the graphs of these two functions, it is clear that an area will form between the two curves. The ratio of this area to the total area $[\theta_{\max} - \theta_{\min}] \times 1$ of the rectangle was defined in [1] as the vagueness, V, of the specialist. That is,

$$V = \frac{1}{2n} \sum_{j=1}^{2n} |\Pi_{\max}(\theta_j) - \Pi_{\min}(\theta_j)| . \qquad (7)$$

In [1] this concept is generalized for the case of multiple experts, and in [2] and [13] this construct is related to other constructs.

Clearly, $0 \leq V \leq 1$. If the specialist answers consistently to all the probability comparability questions, his vagueness will be minimal. It will depend only on the questionnaire itself, that is, on the number and choice of questions. The smaller the number of consistent questions, the greater the vagueness of the specialist. If the expert cannot answer any question at all, his or her vagueness will be one. He or she will be in a state of total ignorance. In this case, the family would consist of all possible probability distributions on θ.

The more vague the specialist is, the less will be its precision, P, and so:

$$P = 1 - V. \tag{8}$$

The constructs defined above could be redefined in local terms. It is necessary just to compute them in a predefined region of the parameter space. It is possible then to search for problematic regions as far as vagueness or precision is concerned. This type of analysis may be useful in identifying some specific inadequacy of the specialist's knowledge or his or her difficulty in expressing it. The region could be just one elementary interval, and then it is possible to establish functions for these constructs, having as arguments the $\theta_j' s$ themselves.

4 Indicators for the Construction of the Elicitation Questionnaire

Two objectives were set for the elicitation questionnaire.

First to permit and guarantee the expert's gradual and smoothly progressive perception of the parameter (state of nature), distributed along the questionnaire. This should be achieved through an increasing of the levels of refinements (decreasing of the coarseness) of the comparability questions so as to avoid retrocession to lower levels already answered. One should try not to confound the specialist with reasoning retrocessions.

Secondly, to permit and guarantee a uniformity of presentation to the expert of the state of nature's elementary intervals distributed along the questionnaire, in such a way as to provide the same choice chance for all ϵ_j. It should be considered here that such uniformity takes into account the fact that the judgment is not based upon preferences but in a degree of belief based on *a priori* knowledge. One tries then not to confuse the expert with a non equitable presentation of elementary intervals. The questionnaire should avoid inducing any bias. It should be neutral. It has to be able to elicite all sorts of distribution.

In setting these two objectives the expert is considered exclusively as a judge facing the alternatives presented to him.

Also, the number of questions should not be too large in order to avoid the fatigue of the expert. Even if the whole interval is divided in a small number of elementary intervals, say, 10 of them, to make all possible comparisons is out of question.

4.1 Question and Elementary Interval Detailing Indicators

To guide the designer in achieving the first objective, a "question detailing indicator", D_{P_i}, is presented. It will guarantee the "local receptivity", i.e., for each question.

For guidance as far as the second objective is concerned, an "elementary interval detailing indicator", D_{ϵ_j}, is presented. It will guarantee the "global receptivity", i.e., for each questionnaire (or part of a questionnaire).

4.2 The Quality of the Questionnaire

First, the symmetry of the questionnaire should be guaranteed. It has to be able to accommodate any distribution shape. If it is not symmetric it could introduce an artificial tail. That is, if the question $P(k, m, s, u)$ was presented to the expert, then the question $P(2n - u, 2n - s, 2n - m, 2n - k)$ must also be presented.

To start the construction of the detailing indicators, a basic detailing indicator, $d_{i,j}$, is used. It is defined by:

$d_{i,j} = 1$ if, in P_i, $\epsilon_j \in I_A \cup I_B$; $d_{i,j} = 0$, otherwise.

The counting indicator for the question will then be defined by

$$d_{P_i} = \sum_j d_{i,j} \tag{9}$$

and the elementary interval counting indicator by

$$d_{\epsilon_j} = \sum_i d_{i,j} \tag{10}$$

d_{P_i} is the number of elementary intervals that appears in question i.

d_{ϵ_j} corresponds to the number of questions in which appears the elementary interval ϵ_j.

The questionnaire can be viewed as a $(q \times 2n)$ matrix whose elements are either 0 or 1.

It is intuitive that the difficulty the expert has in deciding which interval is more likely depends upon the absolute and relative size of these intervals as well as upon the gaps remaining in the overall interval. The problem should be treated then in the two-dimensional space.

Consider then the square

$[\theta_{min}, \theta_{max}] \times [\theta_{min}, \theta_{max}] = [\theta_0, \theta_{2n}] \times [\theta_0, \theta_{2n}]$.

which will have the partition $\wp \times \wp$, with $4n^2$ squared elements with sides measuring $(\theta_j - \theta_{j-1})$. These elements will have as "coordinates" the pairs (x, y), for $x, y = 1, 2, 3, ..., 2n$. The x axis will correspond to the interval I_A and the y axis to the I_B interval. The comparability questions will be then connected unions of elements of $\wp \times \wp$. They will be convex sets (rectangles).

Since $I_A \cap I_B = \emptyset$, and I_A is in the left side of the questionnaire, the feasible questions will correspond to the small squares located in the region delimited by the "coordinates" $(1, 2)$, $(1, 2n)$, and $(2n-1, 2n)$. The squares in the diagonal are forbidden.

The number of elementary squares in this region is given by

$$N = \frac{[(2n-1)+1]\,(2n-1)}{2} = n(2n-1) \qquad (11)$$

These elementary squares work as pixels, defining the "resolution" of the questionnaire, which will be given then by $1/N$. The discernment capability of the expert can be measured also by the size of the smallest elementary interval appearing in a question that he or she can answer consistently with the previous easier ones.

The overall interval $[\theta_0, \theta_{2n}]$ can be normalized so that $[\theta_{2n} - \theta_0] = 1$. Thus, each elementary square will measure $(1/2n) \times (1/2n)$.

The area of the feasible region for constructing questions will be then

$$A_{2n} = \; n(2n-1) \times \frac{1}{(2n)^2} \qquad (12)$$

The area of each question P_i will be given by

$$A_{P_i} = (m-k+1) \times (u-s+1) \times \frac{1}{(2n)^2} \qquad (13)$$

where $1 \leq k \leq m < s \leq u \leq 2n$. The greater the area of the question the easier it will be for the specialist to answer it.

Consider the two particular cases below.

Case 1: $k = 1,\; m = n,\; s = n+1,\; u = 2n$. For this case,

$A_{case1} = \frac{(n-1+1)(2n-\,n-1+1)}{(2n)^2} = \frac{n^2}{4n^2} = 0.25.$

Case 2: $m = k,\; s = u$. For this case,

$A_{case2} = (2n)^{-2} = A_\epsilon.$

The simplest (easiest) question that could be made is

$$P_i = \{\bigcup_{j=k}^{m} \epsilon_j, \bigcup_{j=s}^{u} \epsilon_j\} \qquad (14)$$

where $1 \leq k \leq m < s \leq u \leq 2n$, in the conditions of Case 1. This question compares the two halves of the overall interval. If, in the expert mind, the probability distribution is symmetric, any answer will be good. If it is not symmetric, any expert with a minimum knowledge about the underlying subject will have no difficulty in asserting which one of the halves has more probabilistic mass. The number of squares of this easiest question will be denoted by N_n. This question will be considered as defining the zero of the scale for D_{P_i}.

The less simple question that could be made is the one expressed by (12) in the conditions of Case 2. It compares two elementary intervals. The number of squares of this easiest question will be denoted by N_1. This question will be considered as defining the maximum of the scale, since it will be the most difficult question to answer, and D_{P_i} will have assigned the value 1 corresponding to it.

The question detailing indicator for question P_i will be then defined by:

$$D_{P_i} = \frac{A_{2n}}{A_{case1} - A_{case2}} \left[\left(1 - \frac{A_{P_i}}{A_{2n}}\right) - \left(1 - \frac{A_{case1}}{A_{2n}}\right) \right]$$

$$D_{P_i} = \frac{\frac{n(2n-1)}{(2n)^2}}{\frac{n^2}{(2n)^2} - \frac{1}{(2n)^2}} \left[\frac{\frac{n^2}{(2n)^2}}{\frac{n(2n-1)}{(2n)^2}} - \frac{\frac{(m-k+1)\;(u-s+1)}{(2n)^2}}{\frac{n(2n-1)}{(2n)^2}} \right]$$

$$D_{P_i} = \frac{n^2 \; - (m-k+1)(u-s+1)}{n^2 - 1} \qquad (15)$$

It follows from the definition that $0 \leq D_{P_i} \leq 1$.

Observe also that $(m-k+1) \times (u-s+1)$ is the number of elementary squares representing the question P_i and that no question will have more than n^2 elementary squares.

So, $N_{P_i} = (m-k+1) \times (u-s+1)$, $N_{P_i} \in \{1, 2, ..., n^2\}$, and D_{P_i} will be given by:

$$D_{P_i} = \frac{n^2 \; - \; N_{P_i}}{n^2 - \; 1} \qquad (16)$$

The elementary interval detailing indicator will be defined by:

$$D_{\epsilon_j} = \sum_{i=1}^{q}(d_{i,j} \times D_{P_i}) \qquad (17)$$

If D_{ϵ_j} is such that $\frac{D_{\epsilon_j}}{\sum_{k=1}^{2n} D_{\epsilon_k}} = \frac{D_{\epsilon_l}}{\sum_{k=1}^{2n} D_{\epsilon_k}}$, for all $j, l = 1, 2, 3, ..., 2n$, then the questionnaire is well balanced. All the $D'_{\epsilon_j}s$ will be equal, i.e., the question detailing indicator of all questions will be evenly distributed along all the elementary intervals. If they are normalized to 1 they will constitute a maximum entropy probability distribution over the $2n$ states of nature $\theta'_j s$ (not related to the specialist's knowledge, of course).

5 A Questionnaire Designed using the Question Detailing Indicator

Using a trial and error approach, and following the guidelines suggested above, the questionnaire shown in Table 1 was designed. The total interval (from 0% to 100%) was partitioned in 20 elementary intervals and 42 questions were specified.

Starting from the 50%-50% question, the following questions were constructed in such a way that D_{P_i} would grow roughly linearly with i. A well adjusted regression line of D_{P_i} versus the question number i (1 to 42) has a determination coefficient of 0.98. This can be observed in Table 1.

The $D'_{\epsilon_j}s$ are such that the resulting entropy is equal to 2.9924195, close to the maximum entropy ($log(20) = 2.99573227$).

It is possible to set up a mathematical programming problem in order to automatically design a questionnaire based on those guidelines. As inputs it would have the number of elementary intervals ($2n$) and the desired number of questions. The first question would be the 50%-50% one. The questions should be chosen in such a way as to obtain an as linear as possible evolution of D_{P_i} with i, and an even distribution of the $D'_{\epsilon_j}s$ along the overall interval (sample space).

6 A Practical Application of the Questionnaire

The questionnaire was applied to a medical doctor who is an experienced clinical cardiologist. The following evidence concerning an individual was presented to him: male, aged 49 years, weighting 95 kg, 1.75 m of height, Body Mass Index of $31kg/m^2$,

I_A; I_B	Question Indicator
0-50; 50-100	0.00
0-40; 40-100	0.04
0-60; 60-100	0.04
0-65; 65-100	0.09
0-35; 35-100	0.09
0-70; 70-100	0.16
0-30; 30-100	0.16
0-45; 55-100	0.19
10-50; 50-100	0.20
0-50; 50-90	0.20
0-25; 25-100	0.25
0-75; 75-100	0.25
5-45; 55-100	0.28
0-45; 55-95	0.28
0-40; 60-100	0.36
10-45; 55-95	0.44
5-45; 55-90	0.44
25-75; 75-100	0.51
0-50; 50-75	0.51
25-50; 50-100	0.51
0-25; 50-100	0.51
0-50; 75-100	0.51
0-25; 25-75	0.51
0-45; 50-75	0.56
25-50; 55-100	0.56
0-40; 50-75	0.61
25-50; 60-100	0.61
0-40; 40-60	0.69
40-60; 60-100	0.69
40-80; 80-100	0.69
0-20; 40-80	0.69
25-50; 75-100	0.76
0-25; 25-50	0.76
25-50; 50-75	0.76
0-25; 50-75	0.76
0-25; 75-100	0.76
50-75; 75-100	0.76
30-50; 50-70	0.85
0-10; 70-100	0.89
0-30; 90-100	0.89
20-30; 70-80	0.97
0-10; 90-100	0.97

Table 1: The designed elicitation questionnaire.

US$ 1,300.00 monthly salary, civil engineer, works in a construction firm, has no health complaints, is a nonsmoker, and was randomly chosen amongst all individuals in his city with similar characteristics. The questions concerned the systolic blood pressure of this individual. A minimum of 90 mm Hg and a maximum of 190 mm Hg were established. The questionnaire presented to the cardiologist was the one elaborated above, adding 90 to each entry in the intervals. The following explanation was given in written form to the cardiologist:

> *The questionnaire in annex pertains to the probability distribution of the systolic blood pressure (SBP) of the individual described above. The available evidence, explicited in the ten items above, allows a cardiologist to create an expectation as to what that individual's office SBP will be once it is measured by a sphygmomanometer using the protocol of the British Hypertension Society (average of three measures, etc.). The elicitation protocol aims to translate the knowledge and the experience of the cardiologist in terms of a family of probability distributions for the variable of interest, in this case, the SBP.*
>
> *The items in the questionnaire in annex are all comparisons between the probabilities of the SBP being in one or the other of two intervals. The first question, for instance, is the following: What is more likely, that this individual's SBP is between 90 mm Hg and 140 mm Hg, or that it is between 140 mm Hg and 190 mm Hg? If the cardiologist feels that it is more probable that the SBP is within the first interval, he should mark an "×" on column X1, otherwise, "×" should be marked on column X2.*
>
> *It is important to note that one is not asking what is the interval where the "true" SBP is, for it can actually be in either one or the other. Even with multiple measurements using the sphygmomanometer, or with any other instrument, one will never know for sure. One does not know, but the evidence that is presented allows for an idea of the most probable side.*

He answered all the questions except the last one. He asked for more evidence (family antecedents) in order to feel able to answer it. No more evidence was provided and so he left it blank. His answers were consistent, and the vagueness was 13.75%. Denoting by 1 when he asserted the first interval to be more probable and 0 the other way around, the answers were: 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 1, blank. The graph in Figure 1 shows the results.

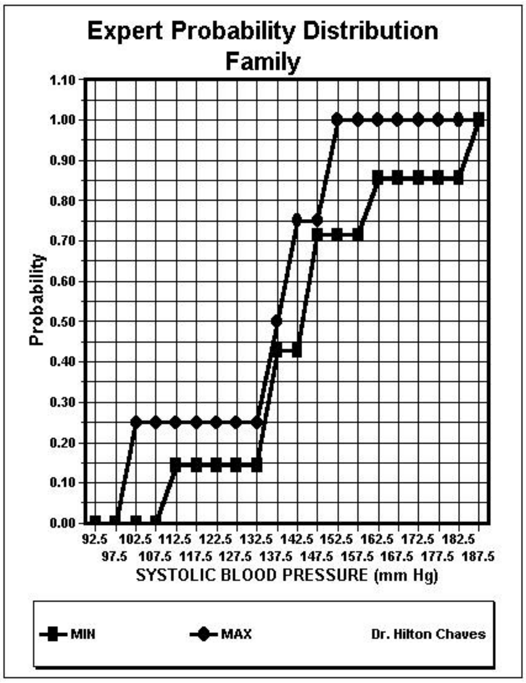

Figure 1: The results of the elicitation.

7 Summary and Conclusions

The method presented here for calculating and representing uncertainty in the epistemic case is compatible with several views of the concept of probability. Betting schemes are avoided, and so the psychological mechanisms of preference do not interfere in the elicitation; the specialist is only asked to make comparisons. It avoids the Bayesian dogma of precision by allowing a range of possible distributions for the parameter. The specialist is not forced to give precise answers. There is no need for statistical or sensitivity analysis.

A question detailing indicator and an elementary interval detailing indicator permitted the elaboration of a questionnaire consisting of paired comparisons of the probabilities of events.

By its very structure the method provides a natural way for the construction of a convex family of probability distributions compatible with the evidence available to the specialist and translated through his answers to the questionnaire. The introduced concepts of vagueness and precision permit a thoroughly evaluation of the elicitation procedure.

The method is practical, and easily implementable.

Acknowledgements

The authors would like to express their gratitude to Alexandre Stamford da Silva, and Luis Cordeiro de Barros Filho, for the helpful discussions and ideas concerning the subject of this article. Stamford elaborated a software that has been very helpful in the development of this research.

References

[1] F. M. Campello de Souza. *Decisoes Racionais em Situacoes de Incerteza*. Tese para concurso de Professor Titular do Departamento de Eletronica e Sistemas da Universidade Federal de Pernambuco, Recife, PE, Brasil, 1993.

[2] F. M. Campello de Souza. Uma Nova Maneira de Calcular e Representar a Incerteza. *Research Report to Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq).* 1997.

[3] G. C. Nadler Lins. *Contribuicoes a um Protocolo de Educao do Conhecimento a Priori*. Tese de Mestrado em Engenharia de Producao, Universidade Federal de Pernambuco, 2000.

[4] James O. Berger, (1985). Statistical Decision Theory and Bayesian Analysis, Second Edition, Springer-Verlag.

[5] Terrence L. Fine, (1973). Theories of Probability: An Examination of Foundations. Academic Press, New York.

[6] Terrence L. Fine, (1977a). An argument for comparative probability. In R. E. Butts and J. Hintikka, editors, Basic Problems in Methodology and Linguistics, pages 105-119. D. Reidel, Dordrecht

[7] Terrence L. Fine, (1977b). Review of A Mathematical Theory of Evidence. Bulletin of the American Mathematical Society, 83:667-672.

[8] Terrence L. Fine, (1983). Foundations of probability. In S. Kotz and N. L. Johnson, editors, The Encyclopedia of Statistical Sciences, volume 3. Wiley, New York (NY).

[9] Terrence L. Fine, (1988). Lower probability models for uncertainty and nondeterministic processes. Journal of Statistical Planning and Inference, 20:389-411.

[10] Glenn Shafer (1976), A Mathematical Theory of Evidence, Princeton University Series.

[11] Peter Walley (1992). Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London, First Edition.

[12] Peter Walley, and Fernando Menezes Campello de Souza, (1990). Uncertainty and indeterminacy in assessing the viability of energy options: a case study of solar heating systems in Brazil, Energy Systems and Policy, Volume 14, pp 281-304.

[13] Fernando Menezes Campello de Souza, (2001). Decisoes Racionais em Situacoes de Incerteza. Book in preparation.