How to take negative information into account?

Samuel Arzola[1], Darien Booth[1], Steve Cruz[1], Miguel Lucero[2], Anaiah Quinn[1], Ana Rodriguez[1], Olga Kosheleva[2], and Vladik Kreinovich[1]

Departments of [1]Computer Science and [2]Teacher Education
University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
searzola@miners.utep.edu, djbooth@miners.utep.edu, sacruz7@miners.utep.edu, malucero3@miners.utep.edu, aequinn@miners.utep.edu, amrodriguez28@miners.utep.edu, olgak@utep.edu, vladik@utep.edu

1. Formulation of the problem

- In many practical situations, we rely on expert estimates.
- What if several experts provide estimates $x_1, \ldots, x_n$ for the quantity of interest?
- Then a natural idea is to use the arithmetic mean $x = (x_1 + \ldots + x_n)/n$ in our decision making.
- But what if some experts provide *negative* information – e.g., saying that the actual value is *not* close to some value y_j?
- How can we take this information into account?
- In this talk, we use decision theory approach to answer this question.

2. What if there is no negative information

- To deal with the problem, let us first consider the case:
  - when there is no negative information, i.e.,
  - when all experts submit some estimates x_i.
- According to decision theory, we should select an alternative with the largest value of expected utility.
- For our situation, this means that we select the value x that maximized the expression

$$p_1 \cdot u_1(x) + \ldots + p_n \cdot u_n(x).$$

- Here, p_i is the probability that the i-th expert is correct, and $u_i(x)$ is the utility according to this expert.
- We often have no reason to assume that some experts are more skilled than others.
- In this case, it is reasonable to consider them equally probable to be correct, i.e., $p_1 = \ldots = p_n$.

3. What if there is no negative information (cont-d)

- What is $u_i(x)$?
- The fact that the i-th expert provides an estimate x_i means that, according to this expert, the utility attains its maximum when $x = x_i$.
- Since they are experts, their estimates x_i are close to the actual value a, so the difference $x_i - a$ is small.
- In this case, we can safely ignore higher order terms in the Taylor expansion of $u_i(x)$ and only keep the first few terms.
- It is not possible to only keep linear terms, since a linear function does not have a maximum.
- So, we need to also take quadratic terms into account, so

$$u_i(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2.$$

- Since the function $u_i(x)$ attains its maximum for $x = x_i$, this means that $u_i(x) = A_i - B_i \cdot (x - x_i)^2$ for some A_i and $B_i > 0$.

4. What if there is no negative information (cont-d)

- There is no reason to assume that the values A_i and B_i are different.
- So, it makes sense to assume that they are the same for all experts, i.e., that $A_i = A$ and $B_i = B$ for some A and B.
- In this case, the expected utility takes the form

$$A - \frac{B}{n} \cdot \sum (x - x_i)^2.$$

- So maximizing it is equivalent to minimizing the sum of the squares

$$\sum (x - x_i)^2.$$

- And it is well known that minimizing this expression leads to the arithmetic mean.

5. What if in addition n positive experts, m experts provide negative information?

- The opinion that the actual value is *not* close to y_j means that the corresponding utility function attains its *smallest* value when $x = y_j$.
- In this case, a similar argument leads to the expression

$$u_j(x) = C_j + D_j \cdot (x - y_j)^2 \text{ for some } C_j \text{ and } D_j > 0.$$

- It still makes sense to assume that the values C_j and D_j are the same for all negative experts: $C_j = C$ and $D_j = D$ for some C and D.
- Thus, the expected utility takes the form

$$\text{const} - \frac{B}{n+m} \cdot \sum (x - x_i)^2 + \frac{D}{n+m} \cdot \sum (x - y_j)^2 /$$

- Let us subtract the constant and divide the maximized function by

$$\frac{B}{n+m}.$$

6. What if we have negative information (cont-d)

- We can conclude that maximizing utility is equivalent to minimizing the sum

$$\sum(x - x_i)^2 - k \cdot \sum(y - y_j)^2, \text{ where } k = D/B.$$

- Differentiating this expression with respect to x and equating the derivative to 0, we get:

$$x = \frac{1}{n - k \cdot m} \cdot (x_1 + \ldots + x_n - k \cdot (y_1 + \ldots + y_m)).$$

7. What if we know that some experts are more reliable

- In this case, instead of assuming that all the probabilities are equal, we know the actual probabilities p_i and q_j.
- In this case, the value x that maximizes the expected utility is

$$x = \frac{1}{P} \cdot (p_1 \cdot x_1 + \ldots + p_n \cdot x_n - k \cdot (q.y_1 + \ldots + q_m \cdot y_m),$$

$$\text{where } P \stackrel{\text{def}}{=} \sum p_i - k \cdot \sum q_j.$$

8. Acknowledgments

This work was supported in part:

- by the US National Science Foundation grants:
  - 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science),
  - HRD-1834620 and HRD-2034030 (CAHSI Includes),
  - EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES),
- by the AT&T Fellowship in Information Technology, and
- by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Focus Program SPP 100+ 2388, Grant Nr. 501624329,