How to Combine Expert Estimates

Christian Gomez, Marcos Lievano, Zachary Wittmann, and Vladik Kreinovich

Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA

cagomez15@miners.utep.edu, malievano1@miners.utep.edu, zawittmann@miners.utep.edu, vladik@utep.edu

Formulation of the problem

- In many practical situations, we rely on experts to estimate the probability of some future event E.
- To get a more accurate estimate, we ask several (n) experts and get their estimates $p_1, \ldots, p_n$.
- It is desirable to come up with a single estimate p that combines these estimates.

Reasonable assumption of independence

- Sometimes, the results of two experts are strongly correlated.
- Then, it makes not much sense to ask both.
- Indeed, the probability provided by the second expert will be practically the same as the probability of the first expert.
- Thus, it makes sense to assume that the experts are independent.

Analysis of the problem

- To analyze this situation, let us consider a more general situation, when:
  - we have n events $E_1, \ldots, E_n$, and
  - expert i estimates the probability of the i-th event E_i.
- In this case, we have 2^n possible situations:

$$E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, E_n,\ E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, \neg E_n,$$
$$E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, \neg E_{n-1} \,\&\, E_n,\ \ldots,$$
$$\neg E_1 \,\&\, \ldots \,\&\, \neg E_{n-1} \,\&\, E_n.$$

- Since the experts are independent, the probability of each situation is equal to the product of the probabilities corresponding to each E_i:

$$p_1 \cdot \ldots \cdot p_{n-1} \cdot p_n,\ p_1 \cdot \ldots \cdot p_{n-1} \cdot (1 - p_n),$$
$$p_1 \cdot \ldots \cdot p_{n-2} \cdot (1 - p_{n-1}) \cdot p_n,\ \ldots,$$
$$(1 - p_1) \cdot \ldots \cdot (1 - p_{n-1}) \cdot (1 - p_n).$$

Analysis of the problem and the resulting formula

- In our case, when there is only one event, we cannot have this event both happening and not happening.
- We have only two possible situations:
  - when E happens – with probability $p_1 \cdot \ldots \cdot p_{n-1} \cdot p_n$.
  - when E does not happen – with probability

$$(1 - p_1) \cdot \ldots \cdot (1 - p_{n-1}) \cdot (1 - p_n).$$

- All other cases are inconsistent.
- Thus, under this condition of consistency, the (conditional) probability p that E will happen is equal to

$$p = \frac{p_1 \cdot \ldots \cdot p_{n-1} \cdot p_n}{p_1 \cdot \ldots \cdot p_{n-1} \cdot p_n + (1 - p_1) \cdot \ldots \cdot (1 - p_{n-1}) \cdot (1 - p_n)}.$$

A natural question: when is the combined probability p larger than an individual estimate p_i – and when is $p < p_i$?

- If we plug in the expression for p into the inequality $p > p_i$ and divide both sides by p_i, we get

$$= \frac{p_1 \cdot \ldots \cdot p_{i-1} \cdot p_i \cdot \ldots \cdot p_{n-1} \cdot p_n}{p_1 \cdot \ldots \cdot p_{n-1} \cdot p_n + (1 - p_1) \cdot \ldots \cdot (1 - p_{n-1}) \cdot (1 - p_n)} > 1.$$

- If we then multiply both sides by the denominator, we get

$$p_1 \cdot \ldots \cdot p_{i-1} \cdot p_i \cdot \ldots \cdot p_{n-1} \cdot p_n >$$
$$p_1 \cdot \ldots \cdot p_{n-1} \cdot p_n + (1 - p_1) \cdot \ldots \cdot (1 - p_{n-1}) \cdot (1 - p_n).$$

- The first term in the right-hand side can be obtained from the left-hand side by multiplying it by p_i.
- Thus, if we move this term to the left0-hand side and take this fact into account, we conclude that

$$(1-p_i) \cdot p_1 \cdot \ldots \cdot p_{i-1} \cdot p_i \cdot \ldots \cdot p_{n-1} \cdot p_n > (1-p_1) \cdot \ldots \cdot (1-p_{n-1}) \cdot (1-p_n).$$

A natural question (cont-d)

- Dividing both sides by $1 - p_i$, we get the following equivalent inequality:

$$p_1 \cdot \ldots \cdot p_{i-1} \cdot p_{i+1} \cdot \ldots p_n >$$
$$(1 - p_1) \cdot \ldots \cdot (1 - p_{i-1}) \cdot (1 - p_{i+1}) \cdot \ldots (1 - p_n).$$

- This is equivalent to $o_1 \cdot \ldots \cdot o_{i-1} \cdot o_{i+1} \cdot \ldots o_n > 1$, where the odds o_j are defined as $o_j \stackrel{\text{def}}{=} p_j/(1 - p_j)$.
- In particular, for $n = 2$ and $i = 1$, this condition is equivalent to $p_2 > 1 - p_2$, i.e., equivalently, to $p_2 > 0.5$.
- This makes perfect sense: $p_2 > 0.5$ means that the second expert is more confident that E will happen than that it will not happen.
- This positive belief in E increase the coverall probability of E.
- Vice versa, if the second expert is more negative about E, this decreases our confidence that E will happen.
- For $n = 3$ and $i = 1$, the above inequality takes the form $p_2 \cdot p_3 > (1 - p_2) \cdot (1 - p_3)$.
- If we open the parentheses and subtract $p_2 \cdot p_3$ from both sides, we get $0 > 1 - p_2 - p_3$.
- This is equivalent to $p_2 + p_3 > 1$ or $(p_2 + p_3)/2 > 0.5$.
- In this case, our confidence increases if, on average, the other two experts believe in E more than in $\neg E$.

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 German Research Foundation).