Structured Expert Judgement

July 2020

Dr Martine J. Barons

With thanks to Dr Anca Hanea
EXPERTS ARE OFTEN CONSULTATION FOR:

- Framing and structuring of a problem
- Identifying variables and relationships
- Sources of data / Data
- Quantifying uncertainty
  - Uncertain parameters (uncertain quantities, probabilities)
Structured expert judgement

**Pre – Elicitation**
- Define problem
- Find experts
- Find facilitator
- Find validation data
- Frame

**Elicitation**

**Post – Elicitation**
- Aggregate experts’ judgements
- Feedback
- Post-hoc analysis of results
WHAT IS THE BEST WAY TO DO IT?

• Nobody knows!
• ...but we know more experts are necessary
• ...and we agree that the following are important:
  • Preparation and planning
  • The selection, phrasing and sequence of questions
  • The aggregation of multiple judgements
  • Documentation of the process
CONDITIONS THAT CHARACTERISE ‘WISE CROWDS’:

- diversity of opinion
- independence
- decentralisation (individuals draw on their own local knowledge)
- aggregation (having a suitable means to generate a group judgement from multiple individual estimates).
• *What* questions we ask the experts, and *how* we ask them, influence their answers

• Psychologists have studied the process of making judgements in uncertainty

• Heuristics and cognitive biases
AGGREGATION

• The process of deriving a single probability (distribution) or a single estimate to represent the knowledge of a group of experts

• Divergent opinions on how best to do this
• Mathematical versus behavioural
An elicitation protocol that:

• asks questions with clear operational meanings
• follows transparent methodological rules
• mitigates psychological and motivational biases
• includes the possibility of identifying the experts
• allows empirical control
• is thoroughly documented
### What % of a watermelon is water?

<table>
<thead>
<tr>
<th>Lower plausible bound</th>
<th>Upper plausible bound</th>
<th>Best estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>95%</td>
<td>50%</td>
</tr>
</tbody>
</table>
Weighting schemes

An expert

Not an expert

Can you spot the difference?
VALIDATION DATA FOR EMPIRICAL CONTROL

- Calibration / performance / seed variables needed
- Scores for good performance
  - Calibration / Statistical accuracy
  - Informativeness
- Use scores as weights
Cooke’s Protocol

- Requires that experts assess uncertainty for variables for which we (will) know the true values:
  
  *Calibration / performance / seed variables*

- Assumption – the future performance of the experts on the variables of interest can be judged on the basis of their past performance on the seed variables
RANGE GRAPHS

calibration questions example

4 experts, 10 calibration questions, correct answer marked.
IDEA (INVESTIGATE, DISCUSS, ESTIMATE, AGGREGATE)

Pre – Elicitation
- Define problem
- Find experts
- Find validation data
- Frame
- Train

Elicitation
- Individual Investigation & 1st set of individual estimates
- Feedback and facilitated Discussion
- 2nd set of individual Estimates

Post – Elicitation
- Aggregate experts’ judgements
- Feedback
- Post-hoc analysis of results
IDEA MINIMISES COGNITIVE BIASES

• The 1st individual assessment avoids anchoring on other people’s estimates

• The discussion between rounds reduces the effect of the availability bias

• The 2nd individual anonymous assessment reduces dominating effects and group think

• The way we ask the questions reduces the anchoring & overconfidence
IDEA PROTOCOL:
WHAT WE’VE LEARNED

• Feedback and facilitated interaction are crucial
• Discussion induces very weak dependence and helps improve experts’ performance
• Aligns expert opinions in the direction of the truth
• Mathematical aggregation is essential
• Equal weighting may be outperformed by unequal weighting
• Performance measures should determine the weights (when unequal)
IDEA:
USE IN Diagram

- Involved archivists from a range of archives to be as general as possible
- Used performance weighted aggregation
- Facilitated discussion between archivists
- Provided data where there was none
- Quantified uncertainty
- Enabled the model to be completed