

ELICITATION RECORD – Part 2

Eliciting a Continuous Distribution

Elicitation title	Risk of widget transmogrification
Workshop	N/A
Date	20 January 2016
Quantity	Population of Portugal (trial run)
Anonymity	In this record, experts are identified by letters A, B, C and the facilitator by Z. All are referred to with male pronouns.
Start time	0925

Definition	The population, in millions, of mainland Portugal as of 2007.
Evidence	For this trial elicitation, none of the participants had specific evidence to record.
Plausible range	Z showed a short presentation to guide the experts in making their individual judgements of the plausible range. Experts wrote their judgements down privately.
Individual elicitation	<p>Method: Tertile</p> <p>Judgements: For each required judgement, Z showed a short presentation to guide the experts in making their individual judgements. Experts wrote their judgements down privately</p>
Fitting	<p>The experts revealed their judgements as follows.</p> <p>Lower plausible bound – A 2, B 5, C 20.</p> <p>Upper plausible bound – A 40, B 20, C 50.</p> <p>Median – A 15, B 10, C 30.</p> <p>Lower tertiles – A 12, B 8, C 27.</p> <p>Upper tertiles – A 20, B 12, C 35.</p> <p>The distributions shown below were fitted using the “Shelf” R package.</p> <p>They were shown to the experts and briefly discussed. The clear disagreement between the experts was noted and Z suggested that this should be discussed fully at the next stage.</p> <p>Looking at the individual distributions, C felt that his fitted distribution indicated much more uncertainty than he actually felt. Z suggested that this should be noted but</p>

that the individual distributions were primarily a reference point for the subsequent exercise of producing an agreed distribution for the group, and in this context it was not important to go back and revise distributions at this time.

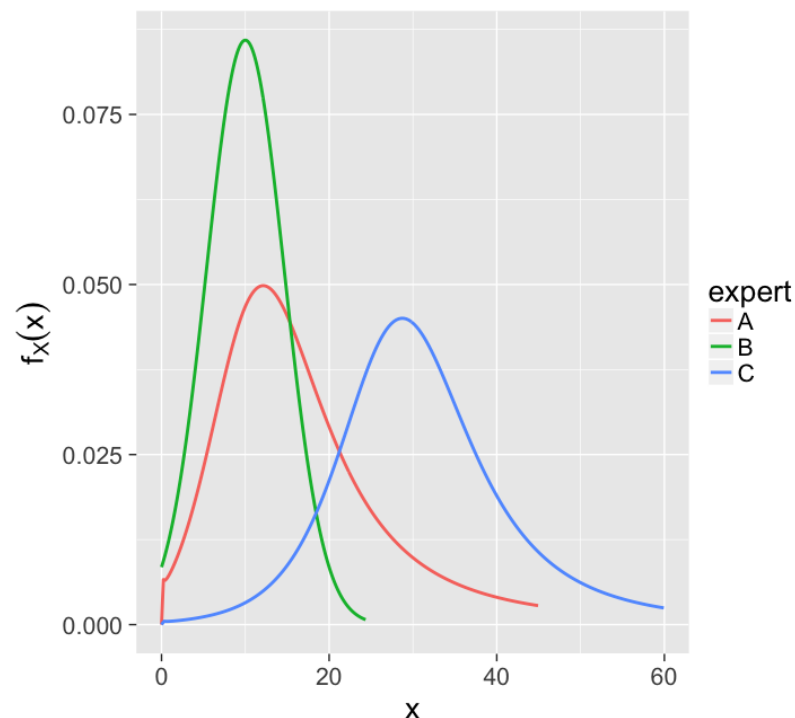


Figure 1: The individual fitted distributions for the population of Portugal (in millions)

Group discussion

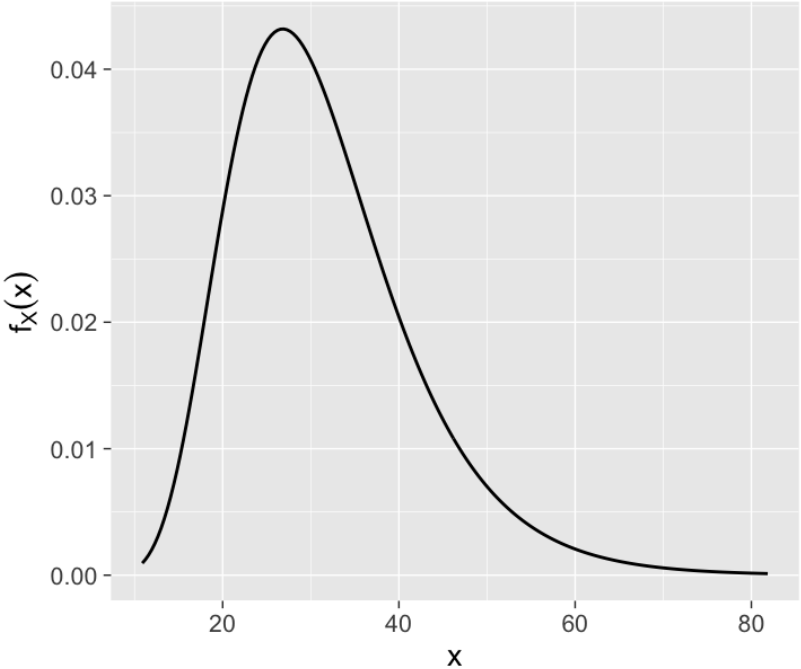
C said that he remembered reading not long ago that the population of Portugal was approximately 30 million. This was the basis for his initial judgements, and he repeated that the fitted distribution reflected too much uncertainty. If he were to repeat the individual elicitation, he would give tighter upper and lower tertiles.

The other participants had given different judgements because they really thought Portugal would have a much smaller population than Britain.

A said that Portugal, while much smaller than Spain, was not a small country (compared for instance to Britain). However, it may be much more sparsely populated than Britain.

C repeated his memory of reading it was 30 million and asked if the others had any such specific evidence to base their opinions on. Z noted that if this were the substantive elicitation rather than just a training exercise, it would have been important to air that information in response to his request for sources of evidence.

A and B were largely persuaded by C's figure.

Group plausible range	<p>Z explained the concept of group consensus judgements in SHELF, and gave a short presentation on the RIO (rational impartial observer) perspective.</p> <p>Z asked the experts to consider the plausible range, noting that in their individual judgements A had thought the population could plausibly be as low as 2, while C thought it might be as high as 50. They agreed on a group consensus range from 5 to 50.</p>
Group elicitation	<p>Method: Probabilities</p> <p>Judgements: After this discussion, Z asked the experts to consider a probability that the population was below 20 million, to represent their combined opinion. They chose 0.1.</p> <p>In the light of this, Z asked first for the probability that the population would be greater than 40 million and second that it was less than 32 million. The group discussed these and agreed on values of 0.2 and 0.6 respectively.</p>
Fitting and feedback	<p>The fitted distribution is shown in the figure below.</p> <p style="text-align: center;">Log normal(3.4, 0.327)</p>  <p style="text-align: center;">x</p> <p>Figure 2: the fitted group distribution for the population of Portugal (in millions)</p> <p>Z gave feedback, first showing that the lognormal distribution fitted their elicited probabilities fairly well. It gave a probability of 0.11 below 20, 0.59 below 32 and 0.18 above 40.</p>

	<p>He also reported that the fitted distribution gave probability of essentially zero for the population being less than 10 million. B said he would have liked a larger probability for this, since his original estimate was 10. However, he acknowledged that the group opinion had changed, being strongly influenced by C's reported figure of 30.</p> <p>Z also noted that the fitted distribution gave a probability of 0.05 that the true population was greater than 50 million. The consensus view was that this was a high probability for what everyone thought was a very unlikely outcome. C, however, noted that if his figure could be too big it could also be too small.</p> <p>The participants agreed to adopt the fitted lognormal distribution as a representation of their combined knowledge about the population of Portugal in 2007.</p>
Chosen distribution	Lognormal(3.4, 0.327), as fitted by "Shelf" software.
Discussion	<p>The participants felt that this had been a really useful exercise, and that they would now approach the substantive elicitation with more confidence. C remarked that he had not realised how close to the median the tertiles should be to reflect sensible amounts of uncertainty. Z responded that while this was a fair comment part of the motivation behind using tertiles was that this factor tended to counter over-confidence.</p> <p>At this point, Z revealed that the true population of Portugal in 2007, according to Portuguese government statistics, was 10.6 million. He suggested that the figure C remembered reading might have been for another country.</p> <p>Z said that he had allowed the group for this training exercise to be strongly influenced by C's claims, but that it showed the risk of over-reacting to data. It was important to weigh the reliability and accuracy of all the evidence.</p>

End time	0950
Attachments	None