

Estadística Bayesiana

Clase 18: Ejemplo Elicitación

Isabel Cristina Ramírez Guevara

Escuela de Estadística
Universidad Nacional de Colombia, Sede Medellín

Medellín, 29 de octubre de 2020

Elicitación-ejemplo

Barrera-Causil C.J, Correa J.C., and Marmolejo-Ramos F. (2019) Experimental Investigation on the Elicitation of Subjective Distributions. *Frontiers in Psychology*. 10: 862.

We examine the resulting personal prior distributions about a percentage when participants receive or do not receive instructions about the elicitation process. Importantly, it is ensured that participants receive the same amount of information about a parameter of interest and a computer application is designed to elicit prior distributions via an interactive questionnaire. This interactive elicitation process provides a distribution of estimates for the parameter of interest for each participant.

Elicitación-ejemplo

Further, a cluster analysis is carried out with the group who received elicitation instructions in order to detect if participants with different degrees of mathematical and/or statistical skills produce distributions of percentages that better capture the parameter of interest (Experiment 1). The elicitation method used in Experiment 1 is then compared with a variation of a graphical elicitation method (Experiment 2). Functional data analysis (FDA) techniques (see Wang et al., 2016) are used to characterize prior distributions of the participants and a novel method is used for clustering distributions (see Methods section for details) (Barrera and Correa, 2015).

Elicitación-ejemplo

EXPERIMENT 1 - Participants. Fifty-nine undergraduate students verbally consented to volunteer for the experiment (agerange = 16–27). Of these participants, 14 had approved a course in basic mathematics and statistics at the university (mathematical and statistical skills group, G1; $\text{Mean}_{age} = 21.7$, $\text{SD} = 2.8$, females = 7), 26 had approved basic mathematics at the university (mathematical skills group, G2; $\text{Mean}_{age} = 20.9$, $\text{SD} = 2.0$, females = 11), and 19 had not completed either basic mathematics or statistics at the university (non-numerical skills group, G3; $\text{Mean}_{age} = 22.8$, $\text{SD} = 2.6$, females = 11).

Elicitación-ejemplo

Participants were pseudo-randomly assigned into the elicitation instruction (I) and non-instruction (NI) groups. Twenty-five participants formed the I group and 34 participants formed the NI group.

Participants in the I and NI groups were informed they would see a random sequence of numbers and images and their task was to determine the percentage of times that the number one appeared (the actual value was 23 %). In order to ensure both groups received the same input information, a fixed random order was used for the presentation of items (phase I). This part of the experiment lasted ~1min. The random sequence of items consisted of 26 items; 10 '1' numbers, 10 '2' numbers, and six images.

Elicitación-ejemplo

Subsequently, both groups of participants underwent the elicitation process (phase II) but only those in the I group received instructions as to what the goal of the elicitation process was. The betting elicitation method was used. This is an interactive method in which the computer application asks questions and provides feedback to participants in order to gauge a range of minimum and maximum estimates and probability values for each. Specifically, the participant is asked about the bets he/she would be willing to place for or against the occurrence of a certain event (E).

Elicitación-ejemplo

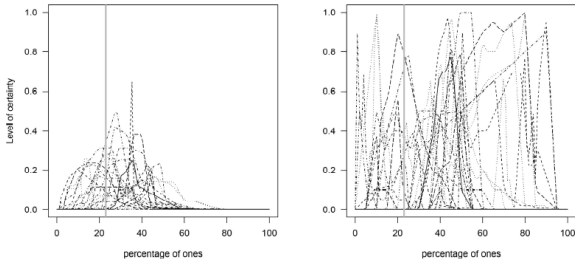


FIGURE 3 | Elicited prior distributions of the percentages given by each of the participants in the I (**Left**) and NI (**Right**) groups. The gray solid vertical line represents the true percentage value (23%).

Elicitación-ejemplo

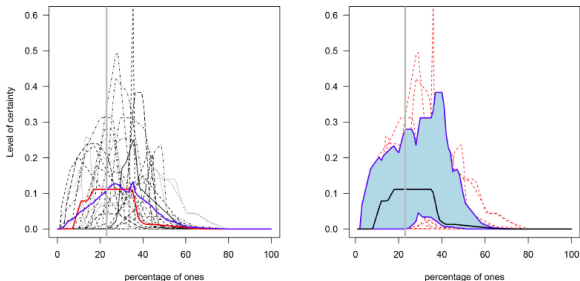


FIGURE 4 | Elicited prior distributions of the percentages in the I group. **Left:** the blue and red lines represent the functional mean and median curves, respectively. **Right:** the solid black line represents the median curve and the lower and upper blue solid lines represent the functional first and third quantiles, respectively (individual distributions are shown in the background in red dotted lines). The gray solid vertical line represents the true percentage value (23%).

Elicitación-ejemplo

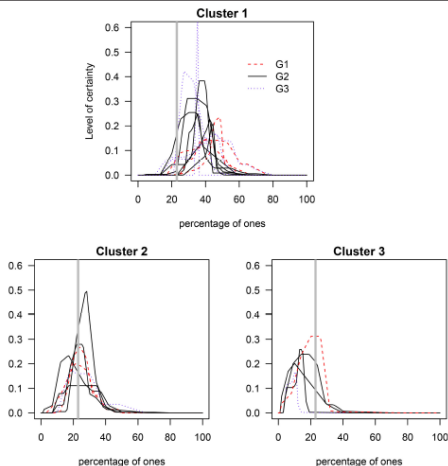


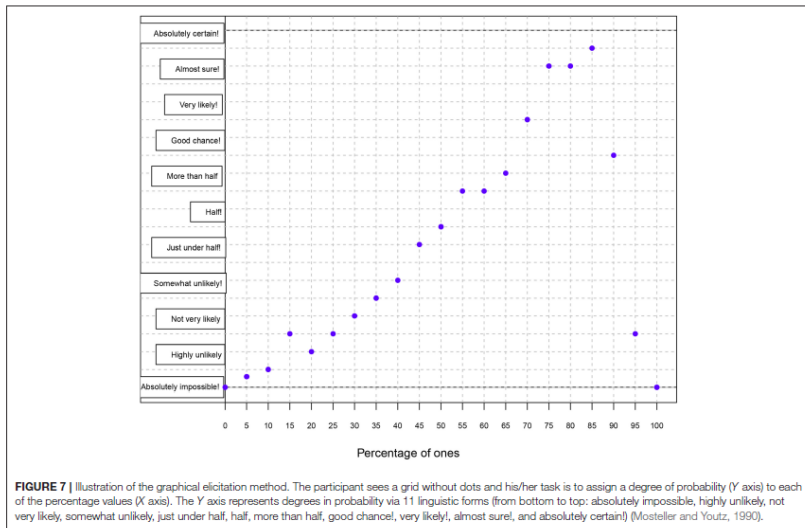
FIGURE 5 | Clusters of the elicited prior distributions of the three groups with varying mathematical and/or statistical skills. G1, mathematical and statistical skills group; G2, mathematical skills group; and G3, non-numerical skills group. The gray solid vertical line represents the true percentage value (23%).

Elicitación-ejemplo

EXPERIMENT 2 - Participants. Thirty-three undergraduate students verbally consented to volunteer in the experiment ($\text{Mean}_{age} = 21.9$, $\text{age}_{range} = 17\text{--}29$, $\text{SD} = 2.5$, females = 16). None of the participants was involved in Experiment 1.

Participants were randomly assigned into two groups: the betting (B) and graphical (G) elicitation groups. The betting elicitation method was the same used in Experiment 1, with the consideration that people were instructed before the elicitation session. The graphical elicitation method enables to represent the degree of knowledge about a parameter of interest via histograms, smooth curves, or points in the Cartesian plane. The ultimate goal is therefore to approximate a probability distribution. In this method, participants are asked to pinpoint on a grid of possible values the level of certainty they have about a parameter. While the X axis represents the values the parameter of interest can obtain, the Y axis represents degrees in probability via adjectives or adverbs of frequency.

Elicitación-ejemplo



Elicitación-ejemplo

Fifteen participants formed the B group ($\text{Mean}_{age} = 21.5$, $age_{range} = 17\text{--}24$, $SD = 2$, females = 8) and 18 participants formed the G group ($\text{Mean}_{age} = 22.3$, $age_{range} = 19\text{--}29$, $SD = 2.9$, females=8). As in Experiment 1, participants in both groups were informed they would see a random sequence of numbers and images and their task was to determine the percentage of times that the number one appeared (the actual value was 77 %). In order to ensure both groups received the same input information, a fixed random order was used for the presentation of items (phase I). This part of the experiment lasted ~ 1 min. The random sequence of items was the same used in Experiment 1. Subsequently, both groups of participants underwent the elicitation process (phase II).

Elicitación-ejemplo

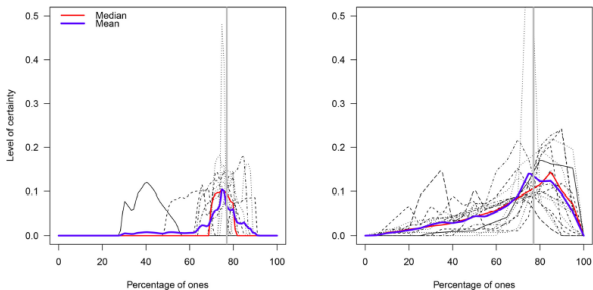


FIGURE 8 | Prior distributions of the percentage of ones in the B and G elicitation groups. The blue and red lines represent the functional mean and median curves, respectively. The gray solid vertical line represents the true percentage value (77%).