MW24.2 Experimental Economics (SS2021)
Probability Judgment

Olexandr Nikolaychuk

1. Bayesian Updating and Representativeness

→ assess the probability of a target event on the basis of:

(i) *base-rate frequency* in the reference population

(ii) *specific evidence* about the case at hand

Taxicab problem [Kahneman and Tversky, 1972a][1]

A cab was involved in a hit and run accident at night. There are two cab companies, *Green* and *Blue*, in the city. You know that:

(a) 85% cabs are *Green* and 15% cabs are *Blue*

(b) a witness, who can correctly identify the color 80% of the time, has identified the cab as *Blue*

(?) what is the probability that the cab was *Blue* rather than *Green*?

$$
\begin{aligned}
P(B|b) &= \frac{P(b|B) \cdot P(B)}{P(b)} = \frac{P(b|B) \cdot P(B)}{P(b|B) \cdot P(B) + P(b|G) \cdot P(G)} \\
&= \frac{0.8 \cdot 0.15}{0.8 \cdot 0.15 + 0.2 \cdot 0.85} = 0.41
\end{aligned}
$$

⇒ median and modal responses are 0.80, which coincides with the witness' credibility (i.e., $P(b|B)$, or *signal precision*) while ignoring the *base rate*, $P(B)$, completely

⇒ omitting (b) makes nearly all subjects respond with 0.15

⇒ omitting (b) and changing '...has identified the cab as *Blue*' into '...has identified the *color* of the cab' results in the median and modal responses of 0.15 to the question: "What is the probability of the witness reporting *Blue*?" (which coincides with the base rate $P(B)$ while in reality, $P(b) = P(b|B) \cdot P(B) + P(b|G) \cdot P(G) = 0.29$)

[1] Daniel Kahneman and Amos Tversky. On prediction and judgment. *Oregon Research Institute Bulletin*, 12(4), 1972a

Alternative formulation:

(a') although the two companies are equal in size, 85% of cab accidents in the city involve *Green* cabs and 15% involve *Blue* cabs

⇒ responses tend to vary a lot; median ∼0.60 (i.e., between 0.80 and 0.41)

---

exa) Casscells et al. [1978][2]:

If a test to detect a disease whose prevalence is $^1/_{1000}$ has a false positive rate of 5%, what is the chance that a person with a positive test result actually has the disease?

⇒ almost half of 60 students and staff at Harvard Medical School report 95%; the average report is 56%; only 11 subjects report 2%

$$P(D|+) = \frac{P(+|D) \cdot P(D)}{P(+|D) \cdot P(D) + P(+|N) \cdot P(N)} = \frac{0.95 \cdot 0.001}{0.95 \cdot 0.001 + 0.05 \cdot 0.999} \approx 0.0187$$

What can explain probability judgements in the taxicab problem?

* Kahneman and Tversky argue that the original formulation leads to the *incidental* interpretation of the *nature* of the base rate whereas the alternative leads to the *causal* interpretation

⇒ *causal nature* (of data) ∼ existence of some *causal* factor is implied that *explains* why any particular instance is more likely to yield one outcome rather than some other

⇒ *incidental nature* (of data) ∼ no such inference is implied

Bar-Hillel [1980][3]:

* version of the taxicab problem where (b) is replaced by a report that the hit-and-run cab was equipped with an intercom while those are installed in 20% of *Green* and 80% of *Blue* cabs

⇒ median response was 0.48

⇒ arguably, the incidental base rate was *not* discarded because the report is "less specific" than the witness' identification (relative to the base rate)

---

[2]Ward Casscells, Arno Schoenberger, and Thomas B. Graboys. Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, 299(18):999–1001, 1978

[3]Maya Bar-Hillel. The base-rate fallacy in probability judgments. *Acta Psychologica*, 44(3): 211–233, 1980

Grether [1980]:

* test of Bayesian learning in an economics laboratory

(?) under-weighing of the base rate as an outcome of the *representativeness* bias

---

Kahneman and Tversky [1972b][4]: "A person who follows this heuristic evaluates the probability of an uncertain event, or a sample, by a degree to which it (i) is similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated"
→ in line with the concept of the *causal* nature of the base rate but *stricter*

[p. 539; excerpt from the original design of Kahneman and Tversky [1973][5]]

* 5 allegedly random vignettes with a personal profile from a sample of 100
* one group was told there were 30 lawyers and 70 engineers in the sample while the other was told there were 70 lawyers and 30 engineers there
* the subjects had to guess to whom a given vignette corresponded
* reminder about the priors after each vignette

⇒ both groups provide nearly identical posterior estimates!

⇒ across the 5 vignettes, the responses are 'correct' in that they follow the intended type of the profile

⇒ 'useless' information ≠ absence of information ('Dick vignette')

Grether's critiques:

* lack of control over *verbal* information
* the vignettes were *not* random, which could be inferred
* the very concept of probability is complicated
* the subjects had to match the 'expert estimates'

[4] Daniel Kahneman and Amos Tversky. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430 – 454, 1972b

[5] Daniel Kahneman and Amos Tversky. On the psychology of prediction. *Psychological Review*, 80(4):237–251, 1973

Grether [1980]: [continued]

* one subject in each session elected to ensure absence of deception etc.
* actually observable random process(es) instead of descriptive data
* sample of 6 draws (with replacement) from either bingo cage A or B, containing $NNNNGG$ or $NNNGGG$ balls, respectively
* varying prior probability of cage A $(\frac{2}{3}, \frac{1}{2}, \frac{1}{3})$
* two repetitions; with and without monetary incentives
* some *combinations* of draws and priors violate *representativeness* [Table I]

Results:

⇒ more mistakes are made (as per Bayes' rule) if representativeness favors the wrong choice [Table III]

⇒ incentives do not seem to matter

⇒ priors are *not* neglected completely [Table IV]

⇒ logit estimates show that the subjects weigh the base rates *less* than the likelihoods $P(sample|cage)$, which supports the use of the representativeness heuristic

⇒ observed samples of $NNNNGG$ or $NNNGGG$ receive extra weight as far as $P(sample|cage)$

⇒ *experience* with a particular sample helps reduce the representativeness bias but does not eliminate it completely

Suggested Literature

- John H. Kagel and Alvin E. Roth, editors. *The Handbook of Experimental Economics*. Princeton University Press, 1995

  [Bayesian Updating and Representativeness; 596–602]

- Daniel Kahneman, Paul Slovic, and Amos. Tversky. *Judgment under uncertainty: heuristics and biases*. Cambridge University Press, Cambridge; New York, 1982

  [Chapter 10. Evidential impact of base rates]

- David M. Grether. Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 95(3):537–557, 1980