14.13 Lectures 7

Xavier Gabaix

February 26, 2004

1 Heuristics and the rules of thumb

- Judgment heuristic: an informal algorithm which generates an approximate answer to a problem.
- Rules of thumb are basically special cases of heuristics.
- Heuristics speed up cognition.
- Heuristics occasionally produce incorrect answers.
- The errors are known as “bias.”
- These are the unintended side effects of generally adaptive processes.

Examples:

- shade your bid in an auction for an oil parcel by 50%
- judge the distance of an object by its clarity
- judge the distance of a person by her size
- save 10% of your income for retirement
- invest (100-age)% of your wealth in stocks

- never borrow on credit cards

- leave a three second interval between you and the car in front of you

- Cognitive psychology studies the representation and processing of information by complex organisms.

- Kahneman and Tversky are two of the leaders in this field.

- They identified three important judgment heuristics in a series of path-breaking contributions in the early 1970's

- representativeness, availability, anchoring

2 Representativeness

Why might similarity poorly predict true probability?

Consider the following example:

"Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations."

Please rank the following statements by their probability, using 1 for the most probably and 8 for the least probable.

1. Linda is a teacher in elementary school.
2. Linda works in a bookstore and takes Yoga classes.
3. Linda is active in the feminist movement.
4. Linda is a psychiatric social worker.
5. Linda is a member of the League of Women Voters.
6. Linda is a bank teller.
7. Linda is an insurance salesperson.
8. Linda is a bank teller and is active in the feminist movement.

1. (5.2) Linda is a teacher in elementary school.
2. (3.3) Linda works in a bookstore and takes Yoga classes.
3. (2.1) Linda is active in the feminist movement.
4. (3.1) Linda is a psychiatric social worker.
5. (5.4) Linda is a member of the League of Women Voters.
6. (6.2) Linda is a bank teller.
7. (6.4) Linda is an insurance salesperson.
8. (4.1) Linda is a bank teller and is active in the feminist movement.

- depending on the subject population, 80%-90% rank item 8 as more likely than item 6.

- K&T call this the conjunction effect (since the conjunctive event receives a HIGHER probability)

- done with naive subjects (undergrads from UBC and Stanford with no background in probability or statistics)

- done with intermediate subjects (graduate students in psychology, education and medicine from Stanford, who had taken several courses in probability and statistics)

- done with sophisticated subjects (graduate students in the decision science program of the Stanford Business School who had taken several advanced courses in probability and statistics)

- results are nearly identical for these three groups

- also similarity ranks perfectly coincide with probability ranks

Bottom line: representativeness heuristic

- Probability follows the conjunction rule: $P(A \cap B) \leq P(B)$.

- The probability that Linda is a feminist bank teller (feminist $\cap$ bank teller $= A \cap B$) is less than the probability that Linda is a bank teller (bank teller $= B$).

- When people face only two choices: Linda is a bank teller or Linda is a bank teller and is active in the feminist movement, they realize that one event includes the other one.

- When those propositions are crowded out by other choices, they use a representativeness heuristic. The description of Linda fits to what a feminist activist could be hence the extra weight associated to this choice.

- Pseudo-Bayesian updating

$$P^{subjective}(\text{activity}_i \mid \text{Description}) \approx kP(\text{Description} \mid \text{activity}_i)$$

Another experiment (conjunction effect: 68%)

Please rank the following events by their probability of occurrence in 1981.

1. (1.5) Reagan will cut federal support to local government.

2. (3.3) Reagan will provide federal support for unwed mothers.

3. (2.7) Reagan will increase the defense budget by less than 5%.

4. (2.9) Reagan will provide federal support for unwed mothers and cut federal support to local governments.

3 Applications of representativeness:

- insensitivity to prior probabilities of outcomes

- insensitivity to sample size

- misconceptions of chance

- insensitivity to predictability

- the illusion of validity?

- misconceptions of regression

3.1 Insensitivity to base rates

- Problem 1:
  - Jack's been drawn from a population which is 30% engineers and 70% lawyers.
  - Jack wears a pocket protector.
  - What is the probability Jack is an engineer?

- Problem 2:
  - Jack's been drawn from a population which is 30% lawyers and 70% engineers.

- Jack wears a pocket protector.
- What is the probability Jack is an engineer?

- We will denote Problem 1 probability by p_1 and Problem 2 probability by p_2

- Consider the events
  - $W = \{$Jack wears a pocket protector$\}$
  - $E = \{$Jack is an engineer$\}$
  - $W = \{$Jack is a lawyer$\}$
- Consider the probabilities
  - $P(E) = q$
  - $P(W/E)$
  - $P(W/L)$

$$
\begin{aligned}
P(E/W) &= \frac{P(W/E)P(E)}{P(W)} \\
&= \frac{P(W/E)P(E)}{P(W/E)P(E) + P(W/L)P(L)} \\
&= \frac{1}{1 + \frac{P(W/L)P(L)}{P(W/E)P(E)}}
\end{aligned}
$$

$$
\Rightarrow \frac{1}{P(E/W)} - 1 = \frac{P(W/L)P(L)}{P(W/E)P(E)} = \frac{P(W/L)}{P(W/E)} \frac{1-q}{q}
$$

- Take the ratio of the probabilities of the two problems

$$
\frac{\frac{1}{p_1} - 1}{\frac{1}{p_2} - 1} = \frac{\frac{1-q_1}{q_1}}{\frac{1-q_2}{q_2}}
$$

$$\Leftrightarrow \frac{\frac{1-p_1}{p_1}}{\frac{1-p_2}{p_2}} = \left(\frac{7}{3}\right)^2 \approx 5$$

- But, in the lab:

$$\frac{\frac{1-p_1}{p_1}}{\frac{1-p_2}{p_2}} \approx 1$$

- Note that $P(W/E)$ and $P(W/L)$ don't need to be evaluated to get the difference between Bayesian updating and the lab results

- What happens when we give the subjects no information other than base rates?

- What happens when we change the description to something uninformative like, "Jack is from Ohio."

3.2 Insensitivity to sample size

- subjects assess the likelihood of a sample result by asking how similar that sample result is to the properties of the population from which the sample was drawn

- A certain town is served by two hospitals.
  - In larger hospital, 45 babies born per day.
  - In smaller hospital, 15 babies born per day.
  - 50% of babies are boys, but the exact percentage varies from day to day. For a period of 1 year, each hospital recorded the days on which more than 60 percent of the babies born were boys.
  - Which hospital do you think recorded more such days?

- 21: the large hospital

- 21: the small hospital

- 53: about the same (within 5% of each other)

- Real life example

  - portfolio not diversified

  - haste to jump to conclusions (about somebody, about a stock)

3.3 Misconceptions of chance (the law of small numbers)

- people expect that a sequence of events generated by a random process will represent the essential characteristics of that process even when the sequence is short

- so if a coin is fair, subjects expect HHH to be followed by a T (gambler's fallacy)

- if girls are as likely as boys, subjects expect GGG to be followed by B

- so BGGBBG is viewed as a much more likely sequence than BBBBBB

- people expect that the essential characteristics of the process will be represented, not only globally in the entire sequence, but also locally in each of its parts

- even scientists make this mistake, overpredicting the likelihood that small sample results will replicate on larger samples

- All families of six children in a city were surveyed. In 72 families the exact order of births of boys and girls was GBGBBG.

- What is your estimate of the number of families surveyed in which the exact order of births was BGBBBB?

In standard subject pools $\approx$20% get it right and the median estimate is 30.

3.4 Misconceptions of regression to the mean

- extreme outliers tend to regress toward the mean in subsequent trials (e.g., best performers on the midterm, fighter pilots with the best landings, tall fathers)

- but intuitively, we expect subsequent trials to be representative of the previous trial, so we fail to anticipate regression to the mean

4 Availability

people assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind

- What percentage of commercial flights crash per year?

- What percentage of American households have less than $1,000 in net financial assets, including savings accounts, checking accounts, CD's, stocks, bonds, etc... (but not counting their most recent paycheck or their defined benefit and defined contribution pension assets)?

5 Anchoring

Anchors seem to matter:

E.g., starting points, frames, defaults, etc....

- Is the Mississippi River more or less than 70 miles long? How long is it?

- Is the Mississippi River more or less than 2000 miles long? How long is it?

Mississippi (mi)	70	2000	300	1500
Everest (ft)	2000	45500	8000	42550
Meat (lbs/year)	50	1000	100	500
SF to NY (mi)	1500	6000	2600	4000
Tallest Redwood (ft)	65	550	100	400
UN Members	14	127	26	100
Female Berkeley Profs	25	130	50	95
Chicago Population (mil.)	0.2	5.0	0.6	5.05
Telephone Invented	1850	1920	1870	1900
US Babies Born (per day)	100	50000	1000	40000

My feelings: Anchoring effects are strongest when anchors have implicit information value and when subjects don't have much time to think about the problem.

Kahneman and Tversky's first anchoring experiment:

- subjects were asked to estimate the percentage of African countries in the UN
- first spin Wheel of Fortune $\rightarrow$ random number
- guess whether % African $>$ random number
- then guess % African
- spin $= 10 \rightarrow$ % African $= 25$
- spin $= 60 \rightarrow$ % African $= 45$

Similar experiment run in class.

- What day of the month were you born?
- Call this number x and write it down
- Let $z =$ % African countries in UN
- What is the value of z?

- What month were you born?
- Call this number x.
- $y = 3 * x$
- write down y
- Let $z=$ % South-American countries in UN
- Is y larger or smaller than z
- What is the value of z?

- Anchoring between the birth day and the percentage of African countries in the UN (t-stat under the coefficients, 20 obs, $R^2 = .23$)

$$\# \, African \;=\; 25.74 - 0.56 \; day$$
$$(6.51) \quad (2.35)$$

- Anchoring between three times the birth month and the percentage of South American countries in the UN (t-stat under the coefficients, 15 obs, $R^2 = .35$)

$$\# \, South \; American \;=\; 7.26 + 0.20 \; 3month$$
$$(4.85) \; (2.69)$$