Lecture 7: Bayesian Information Processing

EC 404: Behavioral Economics
Professor: Ben Bushong

December 7, 2021

Errors in Information Processing

Many interesting questions in economics involve making *inferences* — when you receive some information about something/someone, you must process that information to infer something about the underlying type.

Examples:

- After observing a stock's recent performance, you must decide whether it is a good vs. bad stock to invest in.
- After taking a test drive, you must decide whether you have found a good vs. bad car.
- After interviewing a person, you must decide whether she is a good vs. bad job candidate.
- After receiving a positive medical-test result, you must assess the likelihood that you have the disease.

Linda the Bank Teller

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. Which is more probable?

A : Linda is a bank teller.

B : Linda is a bank teller and is active in the feminist movement.

The Standard Model: Bayesian Information Processing

I'll introduce the model in four pieces:

- Piece 1: Defining types and outcomes.
- Piece 2: Forecasts given uncertainty over underlying types.
- Piece 3: Inferences over types given some information.
- Piece 4: Forecasts given revised beliefs about types.

Piece 1: Defining types and outcomes

Let T denote a set of underlying types.

Let X denote a set of outcomes.

Let $p(x|t)$ denote the probability of outcome $x \in X$ given type $t \in T$.

Example A: Suppose there are two dice, a red one and a blue one. On each die, there are two possible outcomes, H or M. However, on the red die, there are 4 H's and 2 M's, whereas on the blue die, there are 2 H's and 4 M's.

In Example A: $T \equiv \{ \text{ red }, \text{ blue } \}$ $\qquad X \equiv \{ H , M \}$

$$p(H|\text{red}) = 2/3 \qquad p(H|\text{blue}) = 1/3$$

$$p(M|\text{red}) = 1/3 \qquad p(M|\text{blue}) = 2/3$$

Piece 2: Forecasts given uncertainty over underlying types

Let $q(t)$ denote the probability of type $t \in T$ (your beliefs about the distribution of underlying types).

Let $\pi(x)$ denote the probability of outcome $x \in X$ (your forecast for the likelihood of outcome x).

$$\pi(x) \quad = \quad \sum_{t \in T} \; [\; q(t) \; p(x|t) \;] \, .$$

Example B: The underlying types are as in Example A. But suppose I flip a fair coin, and if it comes up heads I'll use the red die, and if it comes up tails I'll use the blue die (and you cannot observe the coin flip). What is the likelihood that the die rolled will come up H?

In Example B: $q(\text{red}) = 1/2$ and $q(\text{blue}) = 1/2$, and so

$$\begin{array}{rclclr} \pi(H) & = & q(\text{red}) \; p(H|\text{red}) & + & q(\text{blue}) \; p(H|\text{blue}) & \\ & = & (1/2)(2/3) & + & (1/2)(1/3) & = 1/2 \end{array}$$

Piece 3: Inferences over types given some information

Let ϕ be a signal (a bit of information).

Let $\gamma(\phi|t)$ be the probability of signal ϕ given type $t \in T$.

Let $q_0(t)$ denote the *prior probability* of type $t \in T$ (your prior beliefs about the distribution of types).

Let $q_1(t|\phi)$ denote the *posterior probability* of type $t \in T$ after having seen signal ϕ (your posterior beliefs about the distribution of types).

Bayes' rule:

$$q_1(t|\phi) = \frac{\gamma(\phi|t)q_0(t)}{\sum_{t' \in T} \gamma(\phi|t')q_0(t')}$$

Piece 3: Inferences over types given some information

Example C: The underlying types are as in Example A, and I flip a coin to pick a die as in Example B. But now I roll that die twice and tell you (truthfully) that the die came up *HH*. What is the likelihood that I'm rolling the red die?

In Example C: $\gamma(HH|\text{red}) = (2/3)(2/3) = 4/9$

$$\gamma(HH|\text{blue}) = (1/3)(1/3) = 1/9$$

Hence:

$$\begin{aligned} q_1(\text{red}|HH) &= \frac{\gamma(HH|\text{red})q_0(\text{red})}{\gamma(HH|\text{red})q_0(\text{red}) + \gamma(HH|\text{blue})q_0(\text{blue})} \\ &= \frac{(4/9)(1/2)}{(4/9)(1/2) + (1/9)(1/2)} = 4/5. \end{aligned}$$

Piece 4: Forecasts given revised beliefs about types

Let $\pi(x|\phi)$ denote the probability of outcome $x \in X$ after having seen signal ϕ (your posterior forecast for the likelihood of outcome x).

$$\pi(x|\phi) \quad = \quad \sum_{t \in T} \left[q_1(t|\phi)\ p(x|t) \right].$$

In Example C:

$$\begin{aligned} \pi(H|HH) \quad &= \quad q_1(\text{red}|HH)\ p(H|\text{red}) + q_1(\text{blue}|HH)\ p(H|\text{blue}) \\ &= \quad (4/5)(2/3) + (1/5)(1/3) = 3/5. \end{aligned}$$

Example: A Medical Test

Suppose there is a rare disease that afflicts 0.01% of the population. You were recently tested for this disease, and your test result was positive. The test is 99% accurate — if you have the disease you will test positive 99 times out of a 100, and if you do not have the disease you will test negative 99 times out of a 100. What is the likelihood that you have the disease?

Note: $q_0(\text{sick}) = 0.0001$ and $q_0(\text{healthy}) = 0.9999$.
$\gamma(+|\text{sick}) = 0.99$ and $\gamma(+|\text{healthy}) = 0.01$.

Hence:

$$\begin{aligned} q_1(\text{sick}|+) &= \frac{\gamma(+|\text{sick})q_0(\text{sick})}{\gamma(+|\text{sick})q_0(\text{sick}) + \gamma(+|\text{healthy})q_0(\text{healthy})} \\ &= \frac{(0.99)(0.0001)}{(0.99)(0.0001) + (0.01)(0.9999)} = 0.0098 = 0.98\%. \end{aligned}$$

Note: If two positive tests: $q_1(\text{sick}|++) = 49.5\%$.

Example: Financial Analyst

Financial analysts predict whether the market will go up or down, and they are sometimes correct and sometimes incorrect. Moreover, there are good analysts and bad analysts. Good analysts are correct 80% of the time, while bad analysts are correct 50% of the time. There are very few good analysts — only 10% of all analysts are good.

Suppose that you have seen your analyst's predictions for the past two months (and not before), and they have both been correct. This month, your analyst is predicting that the market will go down. What should you believe is the likelihood that the market will go down?

Example: Financial Analyst

Step 1: Inference.

$$
\begin{aligned}
q_1(\text{good}|HH) &= \frac{\gamma(HH|\text{good})q_0(\text{good})}{\gamma(HH|\text{good})q_0(\text{good}) + \gamma(HH|\text{bad})q_0(\text{bad})} \\
&= \frac{(0.64)(0.1)}{(0.64)(0.1) + (0.25)(0.9)} \quad = \quad 0.221.
\end{aligned}
$$

Step 2: Forecast.

Since "market will go down" means "your analyst is correct":

$$
\begin{aligned}
\pi(H|HH) &= q_1(\text{good}|HH)p(H|\text{good}) + q_1(\text{bad}|HH)p(H|\text{bad}) \\
&= (0.221) * (0.8) + (0.779) * (0.5) = 0.566.
\end{aligned}
$$

(Counter) Example: Information Cascades

I live in a cul-de-sac in a quiet neighborhood. Every week, I forget which day is garbage day. My neighbors also make the same mistake. We all use each other's behavior to infer garbage day—if you see a bunch of trash cans out, then it's probably time to take out the trash.

Suppose that each of us receives some (private) signal that is accurate $2/3$ of the time, and this is common knowledge. Suppose we put out our trash sequentially.

How often does the neighborhood learn correctly which day is trash day?

Seven (+1) Mistakes in Information Processing

We'll work through seven different mistakes that people make in information processing.

This is not a complete list, and it omits one mistake about other people: *cursedness*.

A person's actions reveal something about their private information (see previous example).

- Selling a used car.
- Buying a stock.
- Offering a product for sale at all.

Question: Do people correctly learn that this very action reveals something about their private information?

If not, we say the person is *cursed* (terminology from Rabin and Eyster).

(1) "Base-Rate Neglect"

"Base-rate neglect":

- People tend to pay too little attention to base rates (priors).

Recall: Bayes' rule is

$$q_1(t|\phi) = \frac{\gamma(\phi|t)q_0(t)}{\sum_{t' \in T} \gamma(\phi|t')q_0(t')}$$

$q_0(t)$ is the base rate or prior.

(1) "Base-Rate Neglect"

Example [Kahneman & Tversky (*ORIRM* 1972)]

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85 percent of the cabs in the city are Green and 15 percent are Blue.

(b) a witness identified the cab as Blue.

The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time. What is the probability that the cab involved in the accident was Blue rather than Green?

(1) "Base-Rate Neglect"

Median answer by subjects was 80%

The correct Bayesian posterior is 41.4%.

$$
\begin{aligned}
q_1(B|b) &= \frac{\gamma(b|B)q_0(B)}{\gamma(b|B)q_0(B) + \gamma(b|G)q_0(G)} \\
&= \frac{(.8)(.15)}{(.8)(.15) + (.2)(.85)} = 0.414
\end{aligned}
$$

(1) "Base-Rate Neglect"

Let $\hat{q}_1(t|\phi)$ denote a person's reported posterior.

Definition: A person exhibits *full base-rate neglect* when her reported posterior $\hat{q}_1(t|\phi)$ is consistent with incorrectly using a uniform prior ($q_0(t)$ is the same for all t), which means

$$\hat{q}_1(t|\phi) = \frac{\gamma(\phi|t)}{\sum_{t' \in T} \gamma(\phi|t')} \equiv q_1^{BRN}(t|\phi).$$

Recall: The correct Bayesian posterior is:

$$q_1(t|\phi) = \frac{\gamma(\phi|t) q_0(t)}{\sum_{t' \in T} \gamma(\phi|t') q_0(t')}$$

Definition: A person exhibits *base-rate neglect* when her reported posterior $\hat{q}_1(t|\phi)$ is in between the correct Bayesian posterior $q_1(t|\phi)$ and the full-base-rate-neglect posterior $q_1^{BRN}(t|\phi)$.

(2) "Law of Small Numbers"

"Law of Small Numbers":

- People tend to exaggerate the degree to which a small sample will resemble the underlying population.

Example [Kahneman &Tversky (*Cog. Psych.* 1972)]

A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be higher than 50%, sometimes lower.

For a period of 1 year, each hospital recorded the days on which more than 60% of the babies born were boys. Which hospital do you think recorded more such days?

Result: 22% answered the larger hospital, 56% answered about the same (within 5% of each other), and only 22% answered the small hospital.

(2) "Law of Small Numbers"

More evidence:

Kahneman & Tversky (*Psych. Bul.* 1971) surveyed mathematical psychologists and found (among other things) that they exaggerated the likelihood that true theories would show up as statistically significant in an experiment with a small number of participants.

$\Rightarrow$ Need to be cautious as scientists to acknowledge how our prior beliefs can shape our analysis of evidence. Might be subconsciously drawn toward thinking we'll find something in the data; theory might be right, but we should be careful.

(3) “Gambler’s Fallacy”

“Gambler’s Fallacy”:

- A belief that recent draws of one outcome increase the odds of the next draw being a different outcome.

“Examples”:

- Roulette: “Black is due.”
- “Lightning never strikes the same place twice.”
- Bias against choosing recent winners in lotteries.

(3) "Gambler's Fallacy"

Terrell (*JRU* 1994) examined New Jersey Pick-3 Lottery.

Each ticket cost 50 cents, and 52% of the money bet is split among winners. Hence, if people pick numbers randomly, then the expected winnings = \$260.

Duration since number last chosen:	Number	Mean Winnings
$<$ 1 week:	8	\$349
1 and 2 weeks:	8	\$349
2 and 3 weeks:	14	\$308
3 and 8 weeks:	59	\$301
$>$ 8 weeks:	1622	\$260
All winners:	1714	\$262

(4) "Over-inference from small samples"

"Over-inference from small samples":

- People tend to infer too much about an underlying probability process from a small sample.

Common procedure [from Grether (*QJE* 1980)]:

- Step 1: Draw a ball from Cage X to determine whether Cage A or Cage B will subsequently be used.
- Step 2: Draw a sample of balls with replacement from that cage.
- Step 3: Show this sample to subjects, and elicit their posterior beliefs.

Common Result: When the sample is more "representative" of one cage than the other, there is over-inference toward that cage.

(4) "Over-inference from small samples"

Recall:

$q_0(t)$ is the person's prior. $q_1(t|\phi)$ is the correct Bayesian posterior. $\hat{q}_1(t|\phi$

Definition: A person exhibits *over-inference* when her reported posterior $\hat{q}_1(t|\phi)$ is such that:

- If $q_1(t|\phi) > q_0(t)$, then $\hat{q}_1(t|\phi) > q_1(t|\phi)$.
- If $q_1(t|\phi) < q_0(t)$, then $\hat{q}_1(t|\phi) < q_1(t|\phi)$.

(5) "Hot-Hand Fallacy"

"Hot-hand fallacy":

- The tendency to perceive positive autocorrelation (a "hot hand") in i.i.d. sequences.

Having hot & cold streaks means something like:

$$\Pr(H|HHH), \Pr(H|HH) > \Pr(H|MM), \Pr(H|MMM).$$

(5) "Hot-Hand Fallacy"

Player-by-player statistics from the 1980-81 76ers:

$\Pr(H\|MMM)$	$\Pr(H\|MM)$	$\Pr(H)$	$\Pr(H\|HH)$	$\Pr(H\|HHH)$
.50	.47	.50	.50	.48
.52	.51	.52	.52	.48
.50	.49	.46	.46	.32
.77	.60	.56	.54	.59
.50	.48	.47	.43	.27
.52	.53	.46	.40	.34
.61	.58	.54	.47	.53
.70	.56	.52	.48	.36
.88	.73	.62	.58	.51

(6) "Conservatism"

"Conservatism":

- People tend to be too conservative in making inferences.

Edwards (1968) finds opposite results from the over-inference studies — that people don't update enough from the samples.

Definition: A person exhibits *under-inference* (or *conservatism*) when her reported posterior $\hat{q}_1(t|\phi)$ is such that:

- If $q_1(t|\phi) > q_0(t)$, then $q_1(t|\phi) > \hat{q}_1(t|\phi) > q_0(t)$.
- If $q_1(t|\phi) < q_0(t)$, then $q_1(t|\phi) < \hat{q}_1(t|\phi) < q_0(t)$.

(6) "Conservatism"

Over-inference from small samples vs. conservatism — how to resolve?

Griffin & Tversky (1992) suggest that people focus too much on the *strength* of evidence and too little on the *weight* of evidence.

- Conservatism studies present people with low-strength/high-weight signals (large samples with relatively even distributions of H vs. M).
- Over-inference studies present people with high-strength/low-weight signals (small samples with skewed distributions of H vs. M).

(7) General lack of understanding of randomness

More generally, people seem to have a lack of understanding of randomness.

- Which birth order is more likely in a family with six children:

$$GBGBBG \qquad \text{vs.} \qquad BGBBBB$$

$$GBGBBG \qquad \text{vs.} \qquad BBBGGG$$

- Suppose there are 23 people in a room — what's the likelihood that there are two people that share a birthday?

  Answer:

An interpretation: "Heuristics and Biases"

Tversky & Kahneman, *Science* 1974

Judgment under uncertainty is a complex task.

People rely on *heuristics* to simplify this task.

In general, these heuristics are quite useful.

But they sometimes lead to systematic and severe *biases*.

An interpretation: "Heuristics and Biases"

Tversky & Kahneman focus on three heuristics:

(1) Representativeness heuristic (similarity): People evaluate the probability of event A by the degree to which it is similar in essential properties to its parent population, and by the degree to which it reflects the prominent features of the process by which it is generated.

(2) Availability heuristic (salience): People evaluate the probability of event A by the ease with which instances and occurrences can be brought to mind.

(3) Anchoring-and-adjustment heuristic: People make judgments by starting from some initial value and then making adjustments, but the adjustments are typically insufficient (Slovic and Lichtenstein, *OBHP* 1971).