Recitation 8: Bayesian Learning

Maddie McKelway

Plan for Recitation

1. Review Chetty et al. (2009) derivation from lecture
2. Bayesian Learning
3. Deviations from Bayesian Learning

Inattention to taxes: Chetty et al. (2009)

- Taxes not featured in price are likely to be ignored
  - Sales tax only added at the register
- Demand $D(\hat{V})$ is a function of perceived value $\hat{V}$
  - Visible part of the value $v = x - p$, where x reflects how much you like the good and p is its price
  - Less visible (opaque) part $o = -tp$, where t is the tax rate
  - $\hat{V} = v + (1-\theta)o = x - p - (1-\theta)\, tp$
  - Note that $\frac{dD}{d\hat{V}} > 0$ (and therefore $\frac{dD}{dp} < 0$)
  - Below focus on opaque part of $\hat{V}$ so write $\hat{V} = v - (1-\theta)\, tp$

Effect of making the tax fully salient

- Would like to compute the change in log demand when θ falls to 0

$$\Delta \log D(\hat{V}) = \log D\left[v - tp\right] - \log D\left[v - (1-\theta)\, tp\right]$$

- Note that for any $f(x)$, $f(x+\alpha) \approx f(x) + \alpha f'(x)$
  - Equivalently, $f(x+\alpha) - f(x) \approx \alpha f'(x)$
  - Let $f(.) = \log D(.)$, $x = v - (1-\theta)\, tp$, and $\alpha = -\theta tp$
  - Then right-hand side above is $f(x+\alpha) - f(x)$, which is $\approx \alpha f'(x)$
- This gives:

$$\begin{aligned} \Delta \log D(\hat{V}) &= \log D\left[v - tp\right] - \log D\left[v - (1-\theta)\, tp\right] \\ &\approx -\theta tp \cdot \frac{d \log D[v - (1-\theta) tp]}{d\theta} \end{aligned}$$

Effect of making the tax fully salient

- Next, note that: $\frac{d \log Y(t)}{dt} = \frac{dY(t)}{Y(t)}$
- This means:

$$\begin{aligned}\Delta \log D(\hat{V}) &= \log D\left[v - tp\right] - \log D\left[v - (1-\theta)\, tp\right] \\ &\approx -\theta tp \cdot \frac{d \log D[v-(1-\theta)tp]}{d\theta} \\ &= -\theta tp * \frac{D'\left[v-(1-\theta)\, tp\right]}{D\left[v-(1-\theta)\, tp\right]}\end{aligned}$$

Effect of making the tax fully salient

- Finally, define the price elasticity of demand $\eta_{D,p}$ as $-\frac{p}{D} \cdot \frac{dD}{dp}$
- This gives:

$$\begin{aligned}
\Delta \log D(\hat{V}) &= \log D\left[v - tp\right] - \log D\left[v - (1-\theta)\, tp\right] \\
&\approx -\theta tp \cdot \frac{d \log D[v - (1-\theta) tp]}{d\theta} \\
&= -\theta tp * \frac{D'\left[v - (1-\theta)\, tp\right]}{D\left[v - (1-\theta)\, tp\right]} \\
&= -\theta t * \eta_{D,p}
\end{aligned}$$

- This implies $\theta = \frac{-\Delta \log D(\hat{V})}{t * \eta_{D,p}}$
  - Chetty et al. (2009) try to measure this

Bayesian Learning: Overview

- Almost all economic decisions are undertaken with some degree of uncertainty
- Individuals must make decisions based on perceived likelihoods of outcomes
- How do individuals form beliefs about statistical likelihoods?
  - Bayesian learning: the “statistically correct” way to form beliefs
  - In reality, we see systematic deviations from Bayesian learning
    - Base rate neglect
    - Gamblers’ fallacy

Bayesian Learning: Overview

- Set-up:
  1. Individual has a *prior* belief of the likelihood that something is true
  2. Individual observes a *signal* in the world that is indicative of whether it's true
  3. Individual combines her prior and signal to form a *posterior* belief of the likelihood that it's true
- How should (in a statistical sense) the individual combine her prior and signal to form a posterior?
  - Use Bayes' Rule!

Bayes' Rule

Notation

- Individual has some hypothesis, h
- Her *prior* belief is that h is true with probability $P(h)$
- She observes *signal*, D, that provides information about the likelihood that h is true
- She forms a *posterior* belief about the probability h is true given D: $P(h|D)$

How does she form $P(h|D)$?

- Bayes' Rule: $P(h|D) = \frac{P(D|h) \cdot P(h)}{P(D)}$
- Where does Bayes' Rule come from?
  - We know $P(h|D) = \frac{P(h \cap D)}{P(D)}$, which implies $P(h \cap D) = P(h|D) \cdot P(D)$
  - Similarly, $P(D|h) = \frac{P(h \cap D)}{P(h)}$, implying $P(h \cap D) = P(D|h) \cdot P(h)$
  - Equating the two expressions for $P(h \cap D)$ gives $P(h|D) \cdot P(D) = P(D|h) \cdot P(h)$ or $P(h|D) = \frac{P(D|h) \cdot P(h)}{P(D)}$

Bayes' Rule Example

Suppose there are two urns: (a) one with an equal number of black and white balls, and (b) one with 75% black balls and 25% white balls. We pick one urn at random and draw a ball at random. The ball drawn in black. What is the probability that we were drawing from urn (a)?

- Should it be greater than, equal to, or less than 0.5?

Notation

- $h =$ ball is from urn (a)
- $D =$ black ball drawn
- We would like $P(h|D) = \frac{P(D|h) \cdot P(h)}{P(D)}$

Bayes' Rule Example

We would like $P(h|D) = \frac{P(D|h) \cdot P(h)}{P(D)}$

- $P(h)$ = the probability the ball comes from urn (a) before we observe the ball's color (the prior probability)
  - What value does $P(h)$ take? 0.5
- $P(D|h)$ = the probability a black ball is drawn if drawing from urn (a)
  - What value does $P(D|h)$ take? 0.5
- $P(D)$ = the probability a black ball is drawn
  - Law of total probability = the probability of an outcome occuring is equal to the sum of probabilities of every distinct way it can occur
  - $P(D) = P(D \cap h) + P(D \cap h') = P(D|h)P(h) + P(D|h')P(h') = (0.50)(0.50) + (0.75)(0.50)$
- Combining: $P(h|D) = \frac{P(D|h) \cdot P(h)}{P(D)} = \frac{(0.5)(0.5)}{(0.50)(0.50)+(0.75)(0.50)} = 0.4$

Base-Rate Neglect

- A common behavioral deviation from Bayesian learning
- *One in a hundred people have HIV, and we have a test for HIV that is 99% accurate. If a person tested positive, what's the probability that she has HIV?*
    - Most people answer 99%
    - Bayes' Rule provides a different answer

Bayes' Rule

- Notation: P=HIV-positive; N=HIV-negative; p=tested positive
- We would like to know $P(P|p) = \frac{P(p|P)P(P)}{P(p)}$
  - $P(p|P)P(P) = (0.99)(0.01)$
  - $P(p) = P(p \cap P) + P(p \cap N) = P(p|P)P(P) + P(p|N)P(N) = (0.99)(0.01) + (0.01)(0.99)$
- This implies: $P(P|p) = \frac{(0.99)(0.01)}{(0.99)(0.01)+(0.01)(0.99)} = 0.5 \neq 0.99$

Base Rate Neglect

- Base Rate Neglect: when given base rate information (i.e. information pertaining to everyone) and specific information (i.e. information pertaining to a particular individual), people focus on the latter and ignore the former
- In the HIV example, people see positive test results (specific information) and forget to account for the fact that HIV is unlikely in the first place (base rate information)
- Implies a deviation from Bayes' Rule

The Gambler's Fallacy

- Another common behavioral deviation from Bayesian learning
- *You toss a coin 20 times. The first 19 times are tails. What's the probability that the final toss is also tails?*
  - Some people might say the probability is very low, reasoning that you've just seen a lot of tails so it would be very unlikely to see another
  - Bayes' Rule gives probability of $\frac{1}{2}$

Bayes' Rule

- Notation: T = the final toss is tails, T_{19} = the first 19 tosses were tails
- Bayes' Rule gives $P(T|T_{19}) = \frac{P(T_{19}|T)P(T)}{P(T_{19})}$
  - Start with $P(T_{19})$, the probability the first 19 draws are tails
    - What value does $P(T_{19})$ take? $\frac{1}{2^{19}}$
  - $P(T_{19}|T)$: the probability that the first 19 draws are tails given the last one is tails
    - Tricky: the last draw being tails tells us nothing about the likelihood that the first 19 were tails
    - The outcomes are independent so $P(T_{19}|T) = P(T_{19}) = \frac{1}{2^{19}}$
  - $P(T)$: probability that the last toss is tails prior to observing the first 19 tosses
    - What value does $P(T)$ take? 0.5
- The $\frac{1}{2^{19}}$'s cancel and we are left with $P(T|T_{19}) = P(T) = 0.5$
  - Intuitively: the signal contains no information so we should stick with our prior
- We didn't have to use Bayes' rule to get this (though going through it is good practice!): could instead have noted that independence means $P(T|T_{19}) = P(T) = 0.5$

The Gambler's Fallacy

- The Gambler's Fallacy: the belief that an event occuring frequently in the past means it's less likely to occur in the future when in fact the occurrences of the event in the past and in the future are independent
- In the coin toss example, many people don't internalize the independence between the last coin toss and the first 19; after they see 19 tails, they think it's very unlikely the 20th would also be tails
- Chen, Moskowitz, & Shue (2016): evidence of the Gambler's Fallacy in high-stakes, real-world decisions
  - Study decisions of asylum judges, loan officers, and baseball umpires
  - Find negative autocorrelation (what does this mean?) of decisions that is unrelated to case quality