CSC200: Lecture 28

Announcements

- Quiz this Friday, February 5
- Quiz is on VCG/GSP for search engine auctions
- Term test Friday, February 12; scope to be announced
- Tomorrows office hours are 12-2 and not 2-4. This term lots of meetings due to faculty and graduate student recruiting.
- Today:
    1. Review and finish up Chapter 16 (information cascades/herding)
    2. Begin Chapter 17 (direct benefit effects)

Spam filtering example from last lecture

- We receive an email with subject heading that more or less says "check this out". Should we treat this as spam?
- Let A be the event that the message is spam.
- Let A' be the event that the message is not spam
- Let B is the event that the message heading is "check this out".
- We are given prior probability $Pr[A] = .4$, $Pr[B|A] = .01$ and $Pr[B|A'] = .004$. We want to compute $Pr[A|B]$.
- By Bayes Rule,

$$Pr[A|B] = Pr[A] \times Pr[B|A]/Pr[B]$$

- We need to calculate $Pr[B]$ and as before:

$$Pr[B|A] \times Pr[A] + Pr[B|A'] \times Pr[A'] = (.01) \times (.4) + (.004) \times (.6) = .0064$$

- Hence $Pr[A|B] = Pr[A] \times Pr[B|A]/Pr[B] = .625$

Additional signals for spam, etc.

- Of course, spam filters (or most uses of Bayes rule) do not rely on only one simple signal (conditional event) such as the occurrence of a particular phrase say in the subject line.

- Rather multiple signals are used to detect spam (e.g. bold face in the subject, occurrence of large sums of money in the email content, etc.).

- These could be called positive signals/events (general or specific to an individual) that raise the probability that a message is spam.

- There may also be negative signals; that is events (usually specific to a given individual) that raise the probability that the message is not spam (e.g. "request to submit assignment late" in subject header).

Bayes with positive and negative signals

- If $B_1, B_2, \ldots, B_r$ are positive signals and $C_1, C_2, \ldots, C_k$ are negative signals then we can just let B (the conditional event) be a (complex) joint event, namely B_1 and B_2 and …and B_r and not C_1 and not C_2 and …and not C_k.
- However, we cannot simply expect that the probability of this joint event is the product of the individual probabilities as it is very likely that these positive and negative signals are not independent.
- With supervised learning, we can attempt to train a spam filter to experimentally determine all the prior and conditional probabilities.

Returning to balls and bins cascade.

- Bayes rule now gives us a quantitative explanation for why a cascade can (and will) form (if the individuals are making decisions implicitly or explicitly on such reasoning).
- The example given made an important assumption (so as to more easily illustrate the cascade or herding phenomena) implying:

> The second person will rely on their own personal information (the ball they see) to break a tie. It will follow that that first two votes honestly report the first two balls drawn.

- So again using the same notation let A be the event that the urn contains two blue balls and one red ball.
- (Since the situation is completely symmetric the same analysis will hold for $\bar{A}$, the event of two red balls and one blue ball in the bin.)
- For the ith individual drawing a ball, we have a different event B depending on the first $i-1$ decisions (different from first $i-1$ balls drawn) and the ith ball drawn.

The first and second individuals

- For the prior probabilities we have $Pr[A] = Pr[\bar{A}] = 1/2$.
- Suppose the first person draws a blue ball. Now the desired conditional probability is $Pr[A|B]$ where B is the event that a blue ball is drawn. Note that the prior probability $Pr[B] = Pr[B|A] \times Pr[A] + Pr[B] \times Pr[\bar{A}] = 1/2$.
- Using Bayes the conditional probability $Pr[A|B]$ is $(1/2) * (2/3)/(1/2) = 2/3$ and the mathematically correct thing to do (i.e. decide or vote what is most probable) is to decide **majority blue**.
- Now suppose the second person draws a red ball. This person can infer that the first person saw a blue ball and now the event B we are conditioning on is the sequence of draws (*blue*, *red*).
- Note we are assuming that each draw is independent (sampling with replacement) but these are not independent events as they both depend on the bin majority and therefore $Pr[B]$ is not simply a product of event probabilities.

Second and third persons

- Using Bayes rule again

$$Pr[A|B] = Pr[A] \times Pr[B|A]/Pr[B] = (1/2) \times (2/9)/(2/9) = 1/2$$

where B is the sequence $(\mathit{blue}, \mathit{red})$ so that

$$Pr[B] = Pr[B|A] \times Pr[A] + Pr[B|\bar{A}] \times Pr[\bar{A}] =$$

$$(1/2)[(2/3) \times (1/3)] + (1/2)[(1/3) \times (2/3)] = 2/9$$

- The second player is now indeed (quantitatively) indifferent and by the tie-breaking assumption of this majority ball example, the second player would decide to say that the bin is **majority red**.
- We are (intuitively) in a situation where the third player can ignore the previous decisions.
- We will instead assume that the second player also saw blue and then will clearly decide blue.
- So now what does the third player decide if he/she draws a red ball?
- Now the event B we are conditioning on is the sequence of draws $(\mathit{blue}, \mathit{blue}, \mathit{red})$ as the third player can be sure that the first two draws were blue.

The cascade has begun

- So far the first two players decided according to the ball that they have drawn.
- In contrast when we calculate Bayes rule for the conditional probability

$$Pr[A|B] = Pr[A] \times Pr[blue, blue, red|A]/Pr[B]$$

- Once again, the denominator cannot simply be calculated by a product of the individual events and a simple calculation shows that $Pr[B] = 1/9$ (and not $1/8$) and hence $Pr[A|B] = 2/3$ so that the third player will ignore their signal and decide blue.
- Further players will infer that only the first two decisions were based on what occurred and will then also act just like the third player and will declare **majority blue**.

Eventually a cascade must begin

- Section 16.6 shows why a cascade will almost surely eventually begin (if there are enough people in this experiment). Let's do this analysis a little differently than the text.
- What we have seen is that the third person will follow their signal if and only if the first two signals are different and otherwise a cascade will form.
- Note that even if the urn is **majority blue**, if the first two balls drawn are red, then the third player will declare the urn is **majority red** and we get a cascade (albeit an incorrect) sequence of decisions.
- Recall, the symmetry in this experiment.

A cascade almost surely forms

- What is the probability that a cascade will not form after say the first $2k$ players have drawn balls?
- This can only happen if each pair $(i, i+1)$ of drawings results in an opposite pair. For definiteness lets say the urn is **majority blue**.
- Then the probability that the first pair are opposites is $2 \times (2/3 \times 1/3) = 4/9$.
- Since we choose to have **majority blue** and **majority red** with equal probability, no matter what is the majority color we have probability $4/9$ to get opposite draws.
- Similarly, the next two colors drawn are opposites with probability $4/9$ and thus the probability that the first k pairs will have opposite colors is $(4/9)^k$ which limits to 0 as k increases.

Lessons from the urn cascade

- Cascades can be wrong and based on little information.
- Cascades can be fragile; e.g. if somehow you see the immediately preceding two (breaking ties as before) or more signals, then the cascade will become that of this latest information.
- Cascades are in contrast to "the wisdom of the crowd" (crowd sourcing) where we hopefully rely on independent observations (or explanations for decisions) before making our own decision.
- The text gives an example of a hiring committee perhaps being subject to a cascade but in reality such decisions are made (let's hope) in a more informative way;
  - Namely, we just don't listen to the votes of previous members but also hear their reasoning. This is analogous to seeing all the balls that have been drawn rather than just hearing all the votes.
  - Or we can vote by closed ballot. More generally, there will be rounds of discussion.

Bayes with payoffs

- Section 16.5 considers a scenario where a sequence of *independent* signals is received and how that would impact a decision when payoffs are involved.

- To do so, the text makes some assumptions to simplify the discussion. The same probabiistic analysis (as for the majority balls experiment) can also be carried over to a (cascading) sequence of decisions when payoffs are involved.

- Let's consider the restaurant example in section 16.5: whether or not to choose a new restaurant over a known restaurant. We are interested in whether or not choosing N (for new) is a good decision G or a bad (or maybe mediocre) decision M. (I am avoiding using B as we used that in the statement of Bayes rule.)

- We have priors $Pr[G] = p$ and $Pr[M] = 1 - p$. (For example, if $p > 1/2$, then it is known that new restaurants tend to be good.)

The payoffs

- The decision: whether or not to choose the new restaurant.
- The payoff/value:
  - for a good decision is $v_g > 0$
  - for a medicore/bad decision is $v_m < 0$.

The payoffs

- The decision: whether or not to choose the new restaurant.
- The payoff/value:
  - for a good decision is $v_g > 0$
  - for a medicore/bad decision is $v_m < 0$.
- A simplifying assumption is that in the absense of additional information, one is indifferent to accepting or rejecting (i.e. staying with the old restaurant) the new restaurant. has payoff 0; namely, we assume

$$(v_g) \cdot p + (v_m) \cdot (1 - p) = 0.$$

  That is, in the absense of any other information, the expected payoff for choosing the new restaurant is the same as staying with the old one.

The payoffs

- The decision: whether or not to choose the new restaurant.
- The payoff/value:
  - for a good decision is $v_g > 0$
  - for a medicore/bad decision is $v_m < 0$.
- A simplifying assumption is that in the absense of additional information, one is indifferent to accepting or rejecting (i.e. staying with the old restaurant) the new restaurant. has payoff 0; namely, we assume

$$(v_g) \cdot p + (v_m) \cdot (1 - p) = 0.$$

  That is, in the absense of any other information, the expected payoff for choosing the new restaurant is the same as staying with the old one.
- We then have access to a set S of independent private signals (e.g. hopefully independent reviews, L and H), for low and high opinions. We assume that these reviews are correct with probability $q > 1/2$. That is, $Pr[L \mid M] = Pr[H \mid G] = q > 1/2$.

When should we choose the new restuarant?

- Suppose we receive one positive signal H. Given this signal, the expected payoff shifts from $0 = v_g Pr[G] + v_b Pr[M] = v_g \cdot p + v_b \cdot (1-p)$ to
$$v_g Pr[G|H] + v_b Pr[M|H] = v_g Pr[G|H] + v_b(1 - Pr[G|H])$$
- Using Bayes rule and the assumption that $q > 1/2$, we have
$$Pr[G|H] = \frac{Pr[G] \cdot Pr[H|G]}{Pr[H]} = \frac{pq}{pq + (1-p)(1-q)} > p$$
That is, the probability of the restaurant being good has increased and hence the expected payoff is now positive and we should try the new restaurant.
- When there is a set S of independent signals where again we assume that these signals tend to be correct with probability $q > 1/2$.
- Not surprisingly, the text shows that if the number a of high signals is greater than the number b of low signals, then the posterior $Pr[G \mid S] > p = Pr[G]$, the prior probability. And for $a < b$ (resp $a = b$) the posterior probability is less than (resp, equal to) p.

Lessons from the urn cascade and the restaurant

- Cascades can be wrong and based on little information
- Cascades can be fragile; if a player gets to see say two (breaking ties as before in favor of private information) or more signals, then the cascade will become that of this latest information.
- Cascades are in contrast to "the wisdom of the crowd" where we rely on what are mainly independent and correct observations (or explanations for decisions) before making our own decision as in the restaurant example.
- As long as the restaurant reviews are indeed independent and tend to be correct, then our probability of making a good choice improves.

Direct benefit effects

- We now turn to Chapter 17 and a study of direct benefit effects
- For example, I might choose a social network (e.g. Linkedin) or an operating system just because its wide spread use gives me a direct benefit (e.g. more people know about me and that might improve my professional standing; more software is available).
- The text refers to this phenomena both as a direct benefit effect and as a "network effect". I think the latter is not the best choice of terminology as neither Chapter 16 nor 17 is in the context of say a social network. Indeed the text says

  > *. . . here, payoffs depend on the number of others who use a good and not on the details of how they are connected.*

- We will discuss social network influence issues later in Chapter 19.

Externalities

- A direct benefit effect is an example of a positive externality. The concept of an **externality** is a little vague but it is an important concept.
- The text explains

  > *An externality is any situation in which the welfare of an individual is affected by the actions of other individuals, without a mutually agreed-upon compensation.*

- Wikipedia gives a "definition" which I find less informative:

  > *In economics, an externality is the cost or benefit that affects a party who did not choose to incur that cost or benefit.*

  The cost of an externality is a negative externality,. . . while the benefit of an externality is a positive externality. . .
- Let's elaborate on what is and what is not an externality by some other examples.

What is and what isn't an externality

- In a combinatorial auction (as we have studied it):
  - Individual items are allocated to different agents.
  - Obviously, another agent valuing an item more than me will affect my utility.
  - But this is not an externality. Why?

What is and what isn't an externality

- In a combinatorial auction (as we have studied it):
  - Individual items are allocated to different agents.
  - Obviously, another agent valuing an item more than me will affect my utility.
  - But this is not an externality. Why?
  - Once I know my allocation (and the price I paid for it), then my utility is determined no matter who gets the other items.

What is and what isn't an externality

- In a combinatorial auction (as we have studied it):
  - Individual items are allocated to different agents.
  - Obviously, another agent valuing an item more than me will affect my utility.
  - But this is not an externality. Why?
  - Once I know my allocation (and the price I paid for it), then my utility is determined no matter who gets the other items.

- On the other hand, let's consider the allocation of ad word slots.
  - We studied this problem as a special case of a combinatorial auction and hence assumed there are no externalities.
  - But it could be (as already mentioned) that my utility for getting a given slot will very much be determined by who gets the other slots.
  - If say competitors get those slots that might very well be a negative externality, whereas if complimentary advertisers got those slots it could be a positive externality.

Markets with a huge number of consumers

The assumption throughout Chapter 17

Any single individual won't affect the aggregate behaviour of the market.

- That is, whether or not I buy a few shares of a particular stock will not impact prices or overall demand.
- But if many people want to buy or sell a given stock then prices will be impacted which in turn will impact further demand.

- A common way to deal with a huge but finite system of individuals is to abstract the system as if individuals are just points on say the real line segment $[0, 1]$. (Apologies for reducing you to a point on the line.)

- Then each individual has no "mass" but subintervals do have proportional mass.

Consumers as points on the line $[0, 1]$

- We assume

Each consumer is a point on the line segment $[0, 1]$ wanting to buy one unit of a good.

- We also assume the consumer's willingness to buy the item depends both on
  1. their intrinsic interest in (i.e. value for) the item and
  2. the number of other people using the good (i.e. the direct benefit effect); the more users the more the item is worth.

- That is, we will be considering a positive externality. What is an example where more "uers" will reduce the worth of an item?

- To start, the text first assumes no direct benefit effect and then studies how direct benefits change things.