Network Science

Lecture 9: 13 August 2009 — Lecturer: Samarth Swarup

Scribe: Chris Kuhlman

Information Cascades

This lecture comes from Chapter 16 of the text.

9.1 Background

This lecture is not about networks per se, although the information can be recast in network terminology. It is more about how crowd mentality develops in a population, in which all folks tend to come to the same conclusion.

9.2 Assumptions of Infomation Cascades

1. Individuals make decisions based on,
   (a) private information not seen by others, and
   (b) actions can be based on actions of previous individuals.
2. No network effects; i.e., no benefit to the individual to act in the same way as others (beyond the immediate decision)
3. All individuals are rational.

Item (1) means that there is some sequencing of decisions made by people: person 1 acts, then person 2 acts (who knows the action taken by person 1), then person 3 acts (who knows the actions taken by persons 1 and 2), etc. For item (2), if a person makes a decision to go to restaurant A for supper, but when he pulls up, the parking lot is empty, but restaurant B, just next door, has many cars in front of it, the person may change their mind to go to restaurant B because many people are there. But there is no extra benefit to the person in changing their decision beyond the immediate desire to have a good tasting meal.

9.3 Example: box with 3 balls

There exists a box with three balls in it. There are two possibilities: (i) there are 2 red balls and 1 blue ball, or (ii) there are 2 blue balls and 1 red ball. Each possibility is equally likely. The first possibility is called *Mr* (for majority red) and the second is called *Mb* (for majority blue). A series of individuals, one at a time, reach into the box, without looking inside, and pull out a ball. They alone look at the ball (i.e., the color of the actual ball drawn from the box is not known to any of the population other than the drawer). The person then makes a public declaration (i.e., the entire population knows the declaration): the person announces whether they think the box is Mr (i.e., the box contains a majority of red balls) or Mb. The table below shows a sequence of the first 4 people of a population. The second column is the private information of drawing a ball; the third column is the composite of what they know: their private information and all preceding public information; and the last column is their public announcement. Each person's declaration is based on the following tie-breaking rule: in case of the same number of red and blue balls in column 3, a person makes the Mr or Mb claim based on the color of the ball that they draw from the box. Remember, all people are rational and there is no misleading the population; everyone is truthful.

Table 9.1: Three balls in a box example where people consecutively and privately select a ball from the box, note its color, and put it back.

person	**color of ball person chooses**	**what they know (private and public info)**	**action (public announcement)**
1	R	R	Mr
2	R	R,R	Mr
3	B	R,R,B	Mr
4	B	R,R,B	Mr

We detail how Table 9.1 is generated. Person 1 sees privately a red ball ("R" in column 2), so her only data point is the private info of "R" in column 3, and so she claims Mr (majority red) in column 4. The second person now knows that person 1 drew a red ball—not because she knows the info in column 2 because that is private. Person 2 knows person 1 drew a red ball because of the public claim in column 4. Say person 2 also draws a red ball ("R"); this is private information in column 2. In column 3, person 2 has the public Mr claim by person 1 (implying an "R") and their own private information of a drawn red ball, so they have 2 "R" (for red balls). So she makes the public announcement of Mr. Person 3 is next; she draws a blue ball, so her private information is "B" and her total information is 2 public "R" and 1 private "B". So, since majority rules, person 3 announces, in column 4, Mr. Now, person 4 is up, and she draws a blue ball, so her private information is "B" in column 2, and her combined private and public information is "R,R,B" in column 3 (the 2 "R" 's are from the public claims of the first 2 people and the "B" is from person 4's own draw. This sequence of 3 values is the same information that person 3 had. Why is not person 4 considering the public announcement of person 3? Because person 4 knows that all people behave rationally, and person 4 knows the rules of the game, so she knows that whether person 3 drew a red or blue ball, they will have either 3 "R"'s or 2 "R"'s and 1 "B"—either way they will announce a majority of red balls; i.e., Mr. R. So person 3's public announcement in column 4 is not relevant to person 4's decision: person 3 provides no new information to person 4. So we have that persons 3 and 4 behave exactly the same. Similarly, all subsequent people will behave the same way, and so we have a cascade: all people from person 3 onward will announce Mr.

More formally, we have:

Definition 1 *An* ***information cascade*** *occurs when people* rationally *choose to follow the crowd.*

By contrast, public voting by union members, by a public show of hands, so that union bosses can monitor their votes, is not an example of an information cascade because individuals are not behaving rationally according to their own choices. There exists intense social pressure to vote in a certain way.

Note also that in the example above, the information cascade described in Table 9.1 could be wrong; e.g., although the cascade of 2 red balls arises, there could in fact be 2 blue balls.

So we have that:

1. information cascades can be wrong.
2. information cascades may be based on very little information.
3. information cascades are fragile: participants must adhere strictly to rules, or else outcomes will change; noise, too, can change the behavior of the information cascade.

9.4 Basic Probability Theory

Let U be a sample space. For example, let X be a random variable corresponding to the outcome of casting a fair 6-sided die. X may be a 1, 2, 3, 4, 5, or 6 from a single roll. An event is specified in terms of these outcomes. Examples: (i) event A is "X is even". This means $A = \{2, 4, 6\}$; (ii) event B is "$X \leq 3$," meaning $B = \{1, 2, 3\}$. The event $A \cap B$ means "X is even and $X \leq 3$;" i.e., $A \cap B = \{2\}$.

The probability of event A is $Pr(A) = |A|/|U| = 3/6 = 1/2$. Also, $Pr(B) = |B|/|U| = 3/6 = 1/2$. $Pr(A \cap B) = |A \cap B|/|A| = 1/6$. The probability of event A, given that event B has occurred is $Pr(A|B) = |A \cap B|/|B| = Pr(A \cap B)/Pr(B) = 1/3$. Similarly, $Pr(B|A) = |A \cap B|/|A| = Pr(A \cap B)/Pr(A) = 1/3$.

The last two sets of equations gives *Bayes' rule*:

$$Pr(B|A)Pr(A) = Pr(A|B)Pr(B). \tag{9.1}$$

Every inference model uses Bayes' rule as a fundamental building block.

Using the following form of Bayes' rule,

$$Pr(A|B) = \frac{Pr(A)Pr(B|A)}{Pr(B)}, \tag{9.2}$$

we say that $Pr(A)$ is the *prior probability* of A and that $Pr(A|B)$ is the *posterior probability* of A.

9.5 Example: Spam email

We want to find the probability that, given an email with the subject line of "check this out," the email message is spam. We will use C for an email with the subject line "check this out." Hence, we want to find $Pr(spam|C)$. We use the last form of Bayes' rule to write:

$$Pr(spam|C) = \frac{Pr(spam)Pr(C|spam)}{Pr(C)}. \tag{9.3}$$

We can compute $Pr(spam) = 0.4$ by examining all email messages over a given period and determining that 40 percent of them are spam. Now, we can again look at previous emails over a given period and determine that 1 percent of the time, spam emails have subject line of "check this out." So, we have $Pr(C|spam) = 0.01$. Similarly, from mining actual emails, we can determine $Pr(C|notspam) = 0.004$. We use these to compute

$$\begin{aligned} Pr(C) &= Pr(C|spam)Pr(spam) + Pr(C|notspam)Pr(notspam) && (9.4) \\ &= (0.01)(0.4) + (0.004)(0.6) = 0.0064, && (9.5) \end{aligned}$$

which is a basic approach for conditional probabilities where "spam" and "not spam" are the conditional events and cover the admissible values of an email message: an email is either spam or not spam. We combine these 3 values to yield

$$Pr(spam|C) = (0.4)(0.01)/(0.0064) = 0.625. \tag{9.6}$$

So note that in order to use Bayes' rule, we need, ideally, large data sets that we can examine to compute input probabilities.

Challenges to using Bayes' rule:

1. estimating probabilities requires lots of data.
2. How to find the correct "features" to evaluate (sometimes this is not obvious, as it was in the example above). An example of a feature is the email subject heading "check this out."
3. identifying feature interdependence.

9.6 Example revisited: box with 3 balls

We initially deduce some probabilities. We have that $Pr(Mb|\ \textit{what was seen and heard}) \geq 1/2$ implies guess Mb (majority blue). Otherwise, guess Mr (majority red). We have $Pr(Mr) = Pr(Mb) = 1/2$ because each of the 2 possibilities is equally likely. Also, we have $Pr(B|Mb) = Pr(R|Mr) = 2/3$ because either way we have 2 balls of the color that we declare to be the majority color, out of 3 trials.

We have the following Table 9.2, where now the blue ball is chosen first. We look at how to calculate a couple of the values in the fourth column.

We compute the probability in the first row: $Pr(Mb|B)$; i.e., given that person 1 sees a blue ball, what is the probability that the majority is truly blue? We compute

$$\begin{aligned} Pr(Mb|B) &= \frac{Pr(Mb)Pr(B|Mb)}{Pr(B)} && (9.7) \\ &= \frac{\frac{1}{2} \times \frac{2}{3}}{\frac{1}{2}} = \frac{2}{3} && (9.8) \end{aligned}$$

Table 9.2: Another example of three balls in a box example where people consecutively and privately select a ball from the box, note its color, and put it back.

person	**color of ball person chooses**	**what they know (private and public info)**	**Calculated** $Pr(Mb\|B)$	**action (public announcement)**
1	B	B	2/3	Mb
2	B	B,B	–	Mb
3	R	B,B,R	2/3	Mb

using direct substitution of values above the table.

Look at the probability for person 3: what is $Pr(Mb|B,B,R)$? Using Bayes' rule, $Pr(Mb|B,B,R) = Pr(Mb)Pr(B,B,R|Mb)/Pr(B,B,R)$. Given that the majority is blue (Mb), a blue ball extraction from the box is 2/3 and a red ball extraction is 1/3. So,

$$Pr(B,B,R|Mb) = \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3} = \frac{4}{27}. \tag{9.9}$$

Now,

$$\begin{aligned} Pr(B,B,R) &= Pr(B,B,R|Mb)Pr(Mb) + Pr(B,B,R|Mr)Pr(Mr) & (9.10)\\ &= \frac{4}{27} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{1}{2} & (9.11)\\ &= \frac{1}{9}. & (9.12) \end{aligned}$$

In this last equation, all conditionals are incorporated: we either have a majority blue (Mb) or a majority red (Mr), and the probabilities from these two contributions are summed. Putting these results together we get, in row 3,

$$\begin{aligned} Pr(Mb|B,B,R) &= \frac{Pr(Mb)Pr(B,B,R|Mb)}{Pr(B,B,R)} & (9.13)\\ &= \frac{\frac{1}{2} \times \frac{4}{27}}{\frac{1}{9}} & (9.14)\\ &= \frac{2}{3}. & (9.15) \end{aligned}$$

The table above shows that an information cascade results with the third person.

9.7 Accept/Reject Model

We now abstract the above discussion into a general model of decision-making behavior. This model is focused on accepting or rejecting some option; e.g., deciding to accept or reject a new technology. Components of it are:

1. Payoff. Each person receives a payoff for their decision. A "reject" decision means payoff is zero. An "accept" decision means payoff is contingent on whether the option is good G or bad

B. If the option is a good idea, then the payoff is $v_g(> 0)$ with probability $Pr(G) = p$, and if the option is a bad idea, then the payoff $v_b(< 0)$ is negative with probability $Pr(B) = (1 - p)$. Here, v_g means a good payoff and v_b means a bad payoff. Assume $v_g p + v_b(1 - p) = 0$, i.e. the *a priori* expected payoff is zero.

2. Signals: they measure private information. H is high; suggests that accepting the option is a good idea. L is low; suggests that accepting the option is a bad idea.

3. States. G is good. B is bad.

We have $Pr(H|G) = q > 1/2$ and $Pr(L|G) = 1 - q < 1/2$. The first equation means that given that an idea is good, high signals are more prevalent than low signals. We assume $Pr(H|B) = 1 - q$ and $Pr(L|B) = q$. See Table 9.3.

Table 9.3: Probabilities for states and signals, where $q > 1/2$.

		State	
		B	G
Signals	L	q	1-q
	H	1-q	q

9.7.1 Decision Making

A priori expected payoff: $v_g Pr(G) + v_b Pr(B) = 0$.

After an H signal: $v_g Pr(G|H) + v_b Pr(B|H)$.

Now

$$Pr(G|H) = \frac{Pr(G)Pr(H|G)}{Pr(H)} \tag{9.16}$$

$$= \frac{pq}{pq + (1 - p)(1 - q)}. \tag{9.17}$$

Because $q > 1/2$, we have $(1 - q) < \frac{1}{2} < q$, so $Pr(G|H) > \frac{pq}{pq+(1-p)q}$, and this reduces to $Pr(G|H) > p$.

Earlier we saw that the third person in the "balls in a box" example uses 3 signals (observations). Let us assume that signals are independent, and that we have a signals that are H and b signals that are L. We have 3 actions depending on the relation between a and b:

1. if $a > b$, accept.
2. if $a < b$, reject.
3. if $a = b$, indifferent.

Consider person N. Let n_{acc} be the number of accepts. Let n_{rej} be the number of rejects. If $n_{acc} - n_{rej} = 0$, then N's signal is the tiebreaker. If $n_{acc} - n_{rej} = 1$, then N's signal (i) makes her indifferent, or (ii) reinforces difference. In all 3 cases, N follows her private signal (because her private signal is followed in case of ties). If $n_{acc} - n_{rej} \geq 2$, then N ignores her signal; this implies a cascade has occurred.

It can be shown that 3 same signals in a row will guarantee a cascade.

As $N \to \infty$, $Pr($ *3 matching signals*$) \to 1$. Also, $Pr($*3 matching signals*$) = q^3 + (1-q)^3$ and $Pr($*no 3 consecutive signals match*$) = [1 - q^3 - (1-q)^3]^{N/3}$.

There are variations on this type of model. Examples include: (1) not all previous decisions are visible (may introduce noise, or network structure to indicate which decisions are observed), and (2) not all signals convey equal information.