Lecture 22: Information Aggregation and The Wisdom of Crowds

Alexander Wolitzky

MIT

6.207/14.15: Networks, Spring 2022

Information Aggregation

Last class, introduced incomplete-information games and auctions.

- Auctions are a way of allocating goods.
- They can also be viewed as a way of **aggregating information**, especially in the case of common value auctions: each bidder has some information about the good, and the bidding in the auction and the final price reflect this information to some extent.

Today, study information aggregation more systematically, focusing on two topics:

1. **Voting:** one-shot collective chocies like elections, jury trials, or corporate decisions.
2. **Prediction markets:** betting markets, where price ends up reflecting participants' information.

Remaining classes: observational learning, social media.

Crowds Make Choices

Many decisions in society are made collectively by groups of individuals with different information.

- What are the prices for shares in various companies on the stock market?
- What are the betting odds for each baseball team to win the World Series?
- Which candidate will win an election?
- Will a defendant by acquited or convicted at trial?
- Which restaurants, books, or movies will become popular?

Wisdom or Folly of the Crowd?

Are crowds wiser or more foolish than the individuals that comprise them?

"It is possible that the many, though not individually good men, yet when they come together may be better, not individually but collectively."
–Aristotle

"No one in this world, so far as I know, has ever lost money by underestimating the intelligence of the great masses of the common people."
–H.L. Mencken

"A large group of diverse individuals will come up with better and more robust forecasts and make more intelligent decisions than even the most skilled decision maker."
–James Surowiecki

"If the blind lead the blind, both shall fall into the ditch."
–Matthew 15:14.

Francis Galton and the Plymouth Ox Weighing Competition

Description from Wikipedia:

“At a 1906 country fair in Plymouth, 800 people participated in a contest to estimate the weight of a slaughtered and dressed ox. Statistician Francis Galton observed that the median guess, 1207 pounds, was accurate within 1% of the true weight of 1198 pounds. This has contributed to the insight in cognitive science that a crowd’s individual judgments can be modeled as a probability distribution of responses with the median centered near the true value of the quantity to be estimated.”

Other Examples

Who Wants to be a Millionaire?

- Accuracy of audience (91%) better than that of expert friend (65%).

Iowa electronic markets.

- First big political prediction market, introduced for 1988 US presidential election.
- Vote share predictions within 1.4% in US presidential elections.
- Better than 75% of opinion polls.

Condorcet

The first person to formally analyze the wisdom of crowds was Marie Jean Antoine Nicolas de Caritat, Marquis de Condorcet, 18th century philosopher, mathematician, politician, collaborator of Leonhard Euler and Ben Franklin, and victim of the French Revolution.

Condorcet made two great contributions to social choice theory.

- **Condorcet's Paradox**: under majority voting among more than 2 alternatives, there can be cycles, where A beats B, B beats C, and C beats A.
  Key forerunner of **Arrow's Impossibility Theorem**, the central result in social choice/voting theory (unfortunately we don't have time to cover it—see EK Ch. 23 if curious).
- The **Condorcet Jury Theorem**: first and most important formalization of the wisdom of crowds.

The Condorcet Jury Theorem

- Imagine a jury, where all jurors have the same preferences: all want to convict the defendant if and only if he's guilty. (Or an election, where all voters have the same preferences and want to elect the best candidate.)
- Each juror/voter has different information: an independent noisy signal of the true state (*guility* or *innocent*).
- Condorcet showed that if all jurors vote according to their information—vote *convict* if your signal indicates guilt, vote *acquit* if your signal indicates innocence—then for a sufficiently large jury, with high probability the majority will vote to convict if the defendant is guilty and vote to acquit if the defendant is innocent.
- As we will see, this is an easy consequence of the law of large numbers (however, Condorcet wrote before the law of large numbers was widely known).
- Galton picked the idea up a century later and developed it into the "wisdom of the crowd."

Condorcet Jury Theorem (cntd.)

Despite its importance, there is a hole in Condorcet's reasoning: if each juror wants *the jury* to convict iff the defendant is guilty, is it actually optimal to *vote* to convict iff your signal indicates guilt?

- Might it ever be optimal to vote **against** your signal?

We will see that the answer is yet, so that developing a fully rational, game-theoretic version of the Condorcet jury theorem is a challenging task.

Plan:

1. Classical, "statistics" version of Condorcet jury theorem, assuming **sincere voting** (i.e., vote your signal).
2. Modern, "game theory" version of Condorcet jury theorem, assuming **strategic voting** (i.e., vote optimally, viewing the election as a game with incomplete information).
3. A bit on prediction markets.

Jury Model

- A defendent is either guilty ($\theta = G$) or innocent ($\theta = I$).
- Prior probability that the defendent is guilty is $p \in (0, 1)$.
  - Everyone shares the same prior belief.
- There are N jurors, who must jointly decide whether to convict ($x = G$) or acquit ($x = I$)
- The jurors have the same preferences: each gets a payoff of
  - 0 if $x = \theta$ (the correct decision is made)
  - $-z$ if $x = G$ but $\theta = I$ (loss for convicting the innocent)
  - $-(1 - z)$ if $x = I$ but $\theta = G$ (loss for acquiting the guilty)

Jury Model (cntd.)

If there is only one juror and she believes the defendant is guilty with probability β, her expected payoff from convicting him is

$$\beta \times 0 + (1-\beta) \times \underbrace{(-z)}_{\text{loss from convicting innocent}},$$

and her expected payoff from acquiting him is

$$\beta \times \underbrace{(-1+z)}_{\text{loss from acquitting guilty}} + (1-\beta) \times 0.$$

Hence, she would convict the defendant if and only if

$$\underbrace{(1-\beta)\,z}_{\substack{\text{expected loss from}\\ \text{wrongful conviction}}} \leq \underbrace{\beta\,(1-z)}_{\substack{\text{expected loss from}\\ \text{wrongful acquittal}}} \iff \beta \geq z.$$

In other words, each juror's preference is that the defendant should be convicted if and only if the probability of guilt exceeds z.

Jury Model (cntd.)

The jurors have different information: each gets a conditionally iid signal $s \in \{G, I\}$ with distribution

$$\Pr(s = G|\theta = G) = \Pr(s = I|\theta = I) = q > 0.5.$$

The decision $x \in \{G, I\}$ is made via a vote among the jurors.

- After observing her signal s_j, each juror j votes $v_j \in \{G, I\}$.
- The decision is $x = G$ if at least k^* voters vote G; it is $x = I$ otherwise.

Leading examples:

- Majority rule: N is odd, $k^* = \frac{N+1}{2}$. (Typical in elections.)
- Unanimity rule: $k^* = N$. (Typical in jury trials.)

Classical CJT

Classical Condorcet Jury Theorem assumes majority rule and **sincere voting**: each juror j "votes her signal", meaning $v_j = s_j$ for each $s_j \in \{G, I\}$.

Theorem

With majority rule and sincere voting, $\Pr(x = \theta)$ *is increasing in* N *and converges to 1 as* $N \to \infty$*.*

- Larger juries are better (the more independent signals, the better).
- Very large juries make the right decision with probability close to 1 (with many signals, very likely that the majority is correct).

Proof (Monotonicity)

Suppose we start with N jurors (N odd) and add 2 new jurors.

- The only way they change the outcome is if the initial vote is 1 short of the correct decision and both new voters vote correctly, or if the inital vote is 1 in favor of the correct decision and both new voters vote incorrectly.
- In either case, we can imagine that the first $N-1$ voters split 50-50, then voter N casts the deciding vote (among the original N jurors), and then the 2 new voters overturn her.
- The probability that voter N gets it right and the 2 new voters get it wrong is $q\left(1-q\right)^2$.
- The probability that voter N gets it wrong and the 2 new voters get it right is $q^2\left(1-q\right)$.
- Since $q > \frac{1}{2}$, we have $q > 1-q$, and hence the latter is more likely.
- Hence, a jury with $N+2$ voters is more likely to make the correct decision than a jury with N voters.

Proof (Convergence)

Identify vote $v = G$ with $+1$ and vote $v = I$ with -1. Let

$$V_N = \sum_{j=1}^{N} v_j$$

- A jury with N voters makes the right decision iff $\theta = G$ and $V_N > 0$, or $\theta = I$ and $V_N < 0$.

Proof (Convergence)

Identify vote $v = G$ with $+1$ and vote $v = I$ with -1. Let

$$V_N = \sum_{j=1}^{N} v_j$$

- A jury with N voters makes the right decision iff $\theta = G$ and $V_N > 0$, or $\theta = I$ and $V_N < 0$.
- Note that

$$E\left[v|\theta = G\right] = q - (1-q) = 2q - 1 > 0,$$

and

$$E\left[v|\theta = I\right] = (1-q) - q = 1 - 2q < 0.$$

Proof (Convergence)

Identify vote $v = G$ with $+1$ and vote $v = I$ with -1. Let

$$V_N = \sum_{j=1}^{N} v_j$$

- A jury with N voters makes the right decision iff $\theta = G$ and $V_N > 0$, or $\theta = I$ and $V_N < 0$.
- Note that

$$E\left[v|\theta = G\right] = q - (1 - q) = 2q - 1 > 0,$$

  and

$$E\left[v|\theta = I\right] = (1 - q) - q = 1 - 2q < 0.$$

- By the weak LLN, for every $\varepsilon > 0$, for each $\theta \in \{G, I\}$,

$$\Pr\left(\frac{V_N}{N} - E\left[v|\theta\right] \; > \varepsilon|\theta \right) \to 0 \text{ as } N \to \infty.$$

  Taking $\varepsilon < 2q - 1$, we see that for each θ, the probability that V_N has the right sign (the group makes the right decision) goes to 1 as $N \to \infty$.

Is Sincere Voting an Equilibrium?

The above argument assumes voters vote sincerely.

- Is this optimal, given that other voters vote sincerely?
- That is, is sincere voting an equilibrium?

Suppose p (prior probability that $\theta = G$) is greater than q (informativeness of one signal), and $z = \frac{1}{2}$ (so optimal to convict iff guilty w/ prob $> \frac{1}{2}$).

Suppose there is only 1 voter, and she gets signal $s = i$. How should she vote?

Is Sincere Voting an Equilibrium?

The above argument assumes voters vote sincerely.

- Is this optimal, given that other voters vote sincerely?
- That is, is sincere voting an equilibrium?

Suppose p (prior probability that $\theta = G$) is greater than q (informativeness of one signal), and $z = \frac{1}{2}$ (so optimal to convict iff guilty w/ prob $> \frac{1}{2}$).

Suppose there is only 1 voter, and she gets signal $s = i$. How should she vote?

Posterior $\Pr(\theta = G | s = I)$ is given by Bayes' rule:

$$
\begin{aligned}
\Pr(\theta = G | s = I) &= \frac{\Pr(\theta = G \cap s = I)}{\Pr(\theta = G \cap s = I) + \Pr(\theta = I \cap s = i)} \\
&= \frac{p(1-q)}{p(1-q) + (1-p)q} \\
&= \frac{1}{1 + \frac{1-p}{p}\frac{q}{1-q}}.
\end{aligned}
$$

Is Sincere Voting an Equilibrium? (cntd.)

$$\Pr\left(\theta = G \middle| s = I\right) = \frac{1}{1 + \frac{1-p}{p}\frac{q}{1-q}}.$$

If $p > q$, this is $> \frac{1}{2}$, so the voter will vote to convict even if her signal indicates innocence.

- The issue is simply that one signal is not enough to overturn the prior.

Is Sincere Voting an Equilibrium? (cntd.)

Now assume there are N voters, and that everyone except voter 1 votes sincerely. How should voter 1 vote?

- She should vote to maximize her expected payoff (and hence everyone's expected payoff, since all voters have the same utility).
- That is, she should vote to maximize $\Pr(x = \theta)$.
- Note that voter 1's vote affects x **only** if the vote among the other $N-1$ voters is a tie. This is the only event in which her vote is **pivotal** for swaying the outcome.
- Posterior belief in this case is given by

$$
\begin{aligned}
& \Pr\left(\theta = G \middle| s = I \cap V_{N-1} = 0\right) \\
= \quad & \frac{p(1-q)\, q^{\frac{N-1}{2}} (1-q)^{\frac{N-1}{2}}}{p(1-q)\, q^{\frac{N-1}{2}} (1-q)^{\frac{N-1}{2}} + (1-p)\, q q^{\frac{N-1}{2}} (1-q)^{\frac{N-1}{2}}} \\
= \quad & \frac{1}{1 + \frac{1-p}{p} \frac{q}{1-q}}.
\end{aligned}
$$

Is Sincere Voting an Equilibrium? (cntd.)

$$\Pr\left(\theta = G \middle| s = I \cap V_{N-1} = 0\right) = \frac{1}{1 + \frac{1-p}{p}\frac{q}{1-q}}.$$

- This is the same as $\Pr\left(\theta = G \middle| s = I\right)$ in the one-voter case where voter 1 is the only voter.
- Just as in the one-voter case , voter 1 will vote to convict even if her signal indicates innocence.
- Sincere voting is **not** an equilibrium.

What is the Equilibrium?

There are always trivial equilibria where the defendant is always convicted or always acquited, regardless of the voters' signals.

- If everyone else always votes to convict, I may as well do so, too.

Let's ignore these, and focus on **symmetric, responsive equilibria**, where all voters use the same strategies and vote to convict with different probabilities for the two possible signal realizations.

Game-Theoretic CJT

Theorem

Suppose $k^*(N) = \alpha N$ for some $\alpha \in (0,1)$ (require at least fraction α guilty votes to convict).
As $N \to \infty$, for any sequence of symmetric responsive equilibria, $\Pr(x = \theta) \to 1$.

- Even though sincere voting is not an equilibrium, the CJT survives in Bayesian Nash equilibrium, for any voting rule other than unanimity.
- We skip the proof, and instead discuss what goes wrong with unanimity.

Unanimity Voting

Unanimity rule is common in jury trials, where a unanimous decision is needed to convict.

Paradoxically, we will see that unanimity rule is **uniquely bad**.

- For any cutoff $\alpha < 1$, innocent defendents are convicted with probability close to 0 with large juries.
- We will see that, with unanimity rule, innocent defendents are convicted with positive probability no matter how large the jury is!
- In particular, innocent defendents are convicted **more** despite a **higher** threshold under unanimity rule, because (as we'll see) rational jurors will vote to convict with high probability even when their signals point to innocence.

Intuition for why Unanimity Voting is Uniquely Bad

- A juror's vote only matters is she is pivotal. Rational juror should condition on the event that they are pivotal.
- Under unamity voting, a juror is pivotal if and only if everyone else votes g. Hence, when casting her vote, a rational juror should assume that everyone else is voting g!
- In a large jury, under since voting, everyone else voting g would be overwhelming evidence of guilt, so a rational juror would ignore her signal and always vote g.
  - Sincere voting is not an equilibrium.
- Moreover, for jurors to rely on their own information **at all**, it must be that everyone else voting g is not overwhelming evidence of guilt.
- For that to be the case, it must be that other jurors use almost the same decision rule regardless of their signals.
  - Note: not true for rules other than unanimity.
- This hinders information aggregation: as $N \to \infty$, $\Pr(x = \theta) \nrightarrow 1$.

What to Make of This?

The model of juries/elections we've considered is very stylized.

- An ingredient in a good understanding of these situations, not the whole story.

The extent to which people actually condition on being pivotal in auctions and elections is controversial (and important!).

- Lab evidence suggests that, given a clear description of situation and time to learn, Bayesian Nash equilibrium can do surprisingly well.
- Battaglini, Morton, Palfrey (2010), "The Swing Voter's Curse in the Laboratory."
- But also plenty of cases where people seemingly fail to condition on pivotality, like the Zillow example.

Much progress made, but how to model and analyze behavior in auctions and elections is still up for debate.

Information Aggregation Beyond Auctions and Elections

So far, we've analyzed two types of "institutions" by which groups aggregate information: **auctions** and **elections**.

A third important example is **markets**.

- Long tradition in economics of viewing market prices as conveying information about the probabilities of different events.
- Goes back at least to a famous article by Friedrich Hayek from 1945, "The Use of Knowledge in Society," arguing that a key advantage of market economies over planned economies is the ability to aggregate information.
- This idea has been very influential in finance, where it takes the form of the **efficient market hypothesis**: the idea that stock prices reflect firms' expected value according to the "market's beliefs."

Prediction Markets

The idea of markets as a vehicle for information aggregation has recently gained influence in the context of **prediction markets**.

- In the simplest prediction market, traders can buy or sell contracts that pay off \$1 if a certain event happens.
  - E.g. as of today, it costs 38¢ to buy a contract that pays off \$1 if Donald Trump is the 2024 Republican presidential nominee.
- It is natural to interpret the price of such contracts as the "market belief".
  - E.g. perhaps the market believes that Trump will be nominated with probability 0.38.
- Interpreting market prices as market beliefs in this way can be useful.
  - Iowa electronic markets outpredicts most opinion polls.
  - Some companies use internal prediction markets to guide business decisions: e.g., some pharmaceutical companies use internal prediction markets to predict what drugs are most likely to pass clinical trials.

Prediction Markets (cntd.)

However, interpreting market prices as market beliefs requires care.

Some issues:

1. The value of \$1 may depend on the realized event.
   - Suppose you think taxes would be higher under President Biden than President Trump. Then \$1 is worth more if Trump is elected, so the market price on Trump will be greater than the "market belief."
   - In general, market prices depend on preferences as well as information. See EK Ch. 22 for details.

2. Market "frictions" may prevent information aggregation.
   - E.g. if prediction market winnings are taxed and losings cannot be easily deducted, you only want to buy/sell if your belief is signficantly different from the price.

Prediction Markets (cntd.)

3. People may want to manipulate the market price, and the total amount of money in the market may be small enough that this is feasible.

- E.g. if you're a Trump supporter, you might want to buy Trump contracts to make him look "strong."
- Rumors of this type of behavior abound in political prediction markets. For example, prediction markets gave Trump a 10% chance to win the 2020 election long after the election was over.

4. Even absent the above issues, if there are limits on betting, price typically does not equal median belief. (Next slide.)

Prediction Markets (cntd.)

Suppose everyone bets the same amount w.
(Betting limit in Iowa Electronic Market is \$500.)

Let $F(\pi)$ be the fraction of people with belief below π, and assume sincere betting: buy iff $\pi > p$, sell if $\pi < p$.
(Questionable assumption, but the simplest possible one.)

If market price is p, the $1 - F(p)$ people with $\pi > p$ buy a total of $(1 - F(p))\, w/p$ contracts; the $F(p)$ people with $\pi < p$ sell a total of $F(p)\, w/(1-p)$ contracts.

Market-clearing price: $1 - F(p) = p$.

Median belief π^m: $1 - F(\pi^m) = 1/2$.

If $\pi^m > 1/2$ then $p \in (1/2, \pi^m)$.
If $\pi^m < 1/2$ then $p \in (\pi^m, 1/2)$.

Intuition: if $\pi^m > 1/2$ then $p > 1/2$, but then optimists afford fewer contracts than pessimists, so p settles between $1/2$ and π^m.

Summary

- The "wisdom of crowds" is a powerful idea that says that collective decisions can be more accurate than individual decisions.
- The classical Condorcet Jury Theorem formalizes this in a model of elections with sincere voting.
- Rational voters do not always vote sincerely, and instead vote optimally conditional on the event that their vote is pivotal.
- With rational voters, sincere voting is not an equilibrium, but information aggregation still occurs in equilibrium for any decision threshold except unanimity rule.
- Prediction and other markets also aggregate information, but the interpetation of market prices as "market beliefs" can be subtle.
- In a simple model of a prediction market with sincere betting, the market price is in between the median belief and 1/2.