Social Networks and Social Choice **Lecture Date: August 31, 2010**

Lecture: Condorcet's Theorem

Lecturer: Elchanan Mossel *Scribes: J. Neeman, N. Truong, and S. Troxler*

Condorcet's theorem, the most basic jury theorem in social choice, is named for Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet (17 September 1743 - 28 March 1794), known as Nicolas de Condorcet. This lecture focuses on the original theorem and some generalizations.

1 The original theorem

Condorcet's Jury theorem applies to the following hypothetical situation: suppose that there is some decision to be made between two alternatives + or −. Assume that one of the two decisions is 'correct,' but we do not know which. Further, suppose there are n individuals in a population, and the population as a whole needs to come to a decision. One reasonable method is a majority vote. So, each individual has a vote X_i, taking the value either $+1$ or -1 in accordance with his or her opinion, and then the group decision is either + or − depending on whether $S_n = \sum_{i=1}^{n} X_i$ is positive or negative.

Theorem 1.1. *(Condorcet's Theorem) [4] If the individual votes $X_i, i = 1, \ldots, n$ are independent of one another, and each voter makes the correct decision with probability $p > \frac{1}{2}$, then as $n \to \infty$, the probability of the group coming to a correct decision by majority vote tends to 1.*

Proof: This is a consequence of the law of large numbers, see [5] Theorem 4.23. Let $a = p - 1/2 > 0$. Since the problem is fair in + and −, we may without loss of generality assume the correct answer is +.

Then $EX_1 = -(\frac{1}{2} - a) + (\frac{1}{2} + a) = 2a > 0$, and the weak law of large numbers states that $\frac{S_n}{n}$ converges in probability to $EX_1 = 2a$, where by converging in probability we mean that for any $\epsilon_1, \epsilon_2 > 0$ there is N large enough such that for every $n \geq N$, $P(|\frac{S_n}{n} - EX_1| < \epsilon_1) > 1 - \epsilon_2$.

Taking $\epsilon_1 = 2a$, we see that the probability of a correct decision is

$$P\left(S_n > 0\right) = P\left(\frac{S_n}{n} > 0\right) \geq P\left(\left|\frac{S_n}{n} - 2a\right| < 2a\right) \to 1,$$

which is what we needed to show. □

This theorem is not realistic for many social situations for a number of reasons, including that individual opinions are not likely to be independent, but for the moment we put such concerns aside. Two natural questions to ask within the context of independence are:

- We know the theorem holds for any fixed $p > \frac{1}{2}$, but what if we let p_n depend on n? How small can p_n be for the theorem to hold?
- The theorem tells us that the probability of a correct decision tends to one as $n \to \infty$, and we will see that for $p_n > \frac{1}{2}$ shrinking slowly enough to $\frac{1}{2}$ it still holds. But what happens for finite n? How large does n need to be for us to know that the probability of a correct decision is high?

- We made our decision based on whether S_n, a function of $X = (X_1, ..., X_n)$, was positive or negative. But what about other aggregation functions? We could, for instance, use $f(X) = X_1$, the dictator function, or the more complicated electoral college function in which individual states take a majority vote, and then we take a weighted majority vote of the states.

The remainder of the lecture will examine these issues.

2 Generalization: p_n varies with n

First, consider the following sequences p_n :

- $p_n = 1 - \frac{1}{n}$.
- $1 > p_n > \frac{1}{2}$ fixed.
- $p_n = \frac{1}{2} + (\log n)^{-1}$
- $p_n = \frac{1}{2} + n^{-1/3}$
- $p_n = \frac{1}{2} + 2^{-n}$.

We will see in this section that the probability that a majority vote gives the correct decision when the individual voters are correct independently with probability p_n tends to one for all of the sequences above except the last. But first, note that for large n the sequences above are nested, i.e. the top one $1 - 1/n$ is eventually larger than any $1 > p > \frac{1}{2}$ which is eventually larger than $\frac{1}{2} + (\log n)^{-1}$, and so on. The following result lets us determine whether the probability of a correct decision tends to one by comparison, instead of having to deal with each case separately:

Proposition 2.1. *Let X_i^n be independent, and equal to 1 with probability p_n and -1 with probability $(1-p_n)$, $1 \leq i \leq n$. Let Y_i^n be independent, and equal to 1 with probability q_n and -1 with probability $(1 - q_n)$, $1 \leq i \leq n$. Define $S_n^X := \sum_i X_i^n$ and $S_n^Y = \sum_i Y_i^n$. Then whenever $p_n \geq q_n$, $P(S_n^X > 0) \geq P(S_n^Y > 0)$.*

Proof: By the binomial formula and the fact that S_n^X is positive iff a majority of the X_i^n are equal to 1, we have

$$P(X_n > 0) = \sum_{n/2 < k \leq n} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = 1 - \sum_{n/2 \geq k} \binom{n}{k} p_n^k (1 - p_n)^{n-k},$$

and

$$P(Y_n > 0) = \sum_{n/2 < k \leq n} \binom{n}{k} q_n^k (1 - q_n)^{n-k} = 1 - \sum_{n/2 \geq k} \binom{n}{k} q_n^k (1 - q_n)^{n-k}.$$

It is enough to show that each term in the sum at right is smaller for p_n than for q_n, and we can do this by proving that for each $k \leq n/2$, $t^k(1-t)^{n-k}$ is decreasing in t for $t > 1/2$. Taking logarithms, we obtain $k \log t + (n - k) \log(1 - t)$, and taking derivatives we have $\frac{k}{t} - \frac{n-k}{1-t}$. Since $n - k \geq k$ and $1 - t \leq t$, the derivative is negative, which is enough. □

Using Proposition 2.1 for the list we made at the start of this section implies that there is a 'cutoff' rate of shrinkage toward $\frac{1}{2}$, above which we get a correct decision in the limit and below which we do not.

To find the cutoff, we appeal to the central limit theorem, which in this context states that if $X_1, X_2, ...$ are i.i.d. taking 1 with probability $p = \frac{1}{2} + a$ and -1 with probability $1/p$, then $\sqrt{n}(S_n/n - 2a)$ has a limiting distribution that is normal with mean and variance equal to those of X_1, i.e. mean $2a$ and variance $4p(1-p)$. By limiting distribution we mean that

$$F_n(t) = P(\sqrt{n}(S_n/n - 2a) \leq t) \to_{n\to\infty} P(Z \leq t)$$

for all t, where $Z \sim N(0, 4p(1-p))$.

If $p_n = a_n + 1/2$ with $a_n \to 0$ then we have that the sum of variances of the first n variables is $n(1+o(1))$. We can apply the Lyapunov central limit theorem (see [2] p. 371), or the central limit theorem for triangular arrays (see [5], Theorem 5.15) to see that the central limit theorem still holds as long as p_n converges, so we obtain that $(S_n - \sum_{j=1}^n a_j)/\sqrt{n}$ converges to a $N(0,1)$. In particular if $a_j << j^{-1/2}$ then $\sum_{j=1}^n a_j/\sqrt{n} \to 0$ and the probability of voting for the correct alternative is approaching 1/2 and if $a_j >> j^{-1/2}$ then $\sum_{j=1}^n a_j/\sqrt{n} \to \infty$ and the probability of voting for the correct alternative is approaching 1.

In particular, returning to our list, all of the sequences given except $\frac{1}{2} + 2^{-n}$ satisfy $\sqrt{n}a_n \to \infty$, so the jury theorem holds for all the sequences except that one.

We can actually see that the majority will not be correct in the limit when $p_n = \frac{1}{2} + 2^{-n}$ even without the central limit theorem: suppose that an individual's opinion is $+$ with probability 1/2, $\pm$ with probability 2^{-n}, and $-$ with probability $\frac{1}{2} - 2^{-n}$. Then the $p = 1/2$ case, in which the jury limit theorem clearly fails because the probability of the majority being correct is exactly $\frac{1}{2}$, corresponds to individuals voting 1 whenever their opinion is $+$ and a vote of -1 whenever it is $\pm$ or $-$. On the other hand, voting with individuals being correct with probability p_n corresponds to a vote of 1 whenever an individual's opinion is $+$ or $\pm$, and a vote of -1 otherwise. But since the expected number of individuals with opinion $\pm$ is $n2^{-n} \to 0$, in the limit there is vanishing chance of any individuals having $\pm$ opinions, hence taking $p_n = \frac{1}{2} + 2^{-n}$ is no better than random voting as $n \to \infty$.

3 Results for finite n

To obtain results regarding the probability of a correct outcome for finite n, we can appeal to large deviations theory. One result by Cramér dating to the early twentieth century [3] states that if $Y_1, Y_2, ..., Y_n$ are iid Bernoulli (p) random variables, i.e. they are 1 with probability p and 0 with probability $1-p$, and $\bar{Y}_n$ denotes their average, then for each $b > 0$,

$$P(|\bar{Y}_n - p| > b) < 2e^{-2b^2 n}.$$

We framed the jury theorem in terms of votes of ± 1 rather than votes of 0 or 1, but if we let $Y_i = 1$ when $X_i = 1$ and $Y_i = 0$ when $X_i = -1$, then the sum of the X_i's is positive iff the average $\bar{Y}_n$ is larger than $\frac{1}{2}$.

Hence, when the probability of a correct vote from each individual is is $p = \frac{1}{2} + a$, then by taking $b = a$ in the large devation result we obtain, using that $t \leq 1/2$ implies $|t - 1/2 - a| \geq a$,

$$P(S_n > 0) = P\left(\bar{Y}_n > \frac{1}{2}\right) \geq P\left(\left|\bar{Y}_n - \frac{1}{2} - a\right| < a\right) > 1 - 2e^{-2a^2 n}.$$

This provides a good bound on the error for finite n, and since nothing here prevents a_n from changing with n since the result holds for each fixed $p = \frac{1}{2} + a$ and n and is not about a limit, this also provides another way of deriving the result from the previous section, that the jury limit theorem holds for $p_n \downarrow \frac{1}{2}$ iff $\sqrt{n}a_n \to \infty$.

4 Electoral College

You will be asked to derive results regarding the actual United States' electoral college as an exercise, but for now consider the following idealized version of the electoral college:

We have $n = m^2$ individuals, grouped into m states with m individuals each. Their opinions are independent and correct with probability p_n. When we vote, each state takes a majority vote from its m members, and then state casts a single vote of ± 1, in agreement with its majority vote, in the larger election.

For fixed p it is easy to see that the jury will still vote correctly, because each state by itself votes correctly with probability tending to one, and then averaging the state votes will give a correct result with probability tending to 1. For variable p_n, the result is less clear. If $a_n\sqrt{m} \to \infty$, then the same reasoning, combined with previous results, shows that the jury will make a correct decision in the limit, but the exact cutoff is not as clear.

It turns out, however, that $\sqrt{n}a_n \to \infty$ (a much weaker condition than $\sqrt{m}a_n \to \infty$) is still enough. This will be part of the homework.

5 Other Aggregation Functions

It is reasonable to suppose that in any society, the decisions made by the overall population will depend in some deterministic way on the opinions held by the individuals. So, we assume that there is some function $f(\cdot)$ such that, when $X \in \{-1, 1\}^n$ describes the individual votes or opinions on an issue, then the society as a whole makes a decision of $+$ whenever $f(X) > 0$ and $-$ whenever $f(X) < 0$.

We would like to explore what the best and worst reasonable choices of f are, i.e. what are the best and worst possible ways of deciding a group opinion from individual opinions, under this model and assuming that the goal is to come to a correct decision as often as possible.

First we need to restrict the possible f to prevent unreasonable choices. First, we assume that f is fair, i.e. $f(-X) = -f(X)$ for all X. We do this because, although, for convenience, we have been letting $+$ denote the correct choice and $-$ the incorrect choice, in the real world we do not know what is correct and incorrect, so the function f has to be fair in order to know which choice to make in realistic situations. For instance $f \equiv +1$, i.e. always making the correct choice, is not reasonable in the real world.

We also assume that f is monotone: i.e. $f(X) \geq f(X')$ whenever $X_i \geq X'_i$ for all f. In words, this says that if the only thing that happens is more people change their opinions to $+$, then the group opinion can only change toward $+$. This ensures that our 'worst' choice will not be something stupid, like the minority opinion.

It turns out that these two simple restrictions are enough to determine the best and worst possible functions f. The best is the majority vote $f(X) = \sum_i X_i$, and the worst is the dictator function $f(X) = X_1$. The proof that the dictator function is worst is a bit involved and will be postponed until the next lecture, but we can prove easily that the majority vote is best:

Proposition 5.1. *The best fair monotone function f, in the sense that when the true state of the world is $\pm$ then $\text{sign}(f(X))$ is $\pm$ with maximal probability and $\mp$ with minimal probability, is the majority vote function.*

Proof: Because we have restricted attention to fair f, it is enough to find the Bayes procedure [1] under the prior that the world's true state is $\pm$ with probability one-half each, since by symmetry the probability of a correct decision is the same regardless of the state of the world, and hence the Bayes procedure is also the best under any particular state.

So, let the true state S of the world equal 1 with probability 1/2 and -1 with probability 1/2, and let $P(X_i = s|S = s) = p > 1/2$ and $P(X_i = -s|S = s) = 1 - p$. Then we can use Bayes rule to compute the probabilities

$$P(S = 1|X) = \frac{P(X|S = 1) \cdot \frac{1}{2}}{P(X)}, \qquad P(S = -1|X) = \frac{P(X|S = -1) \cdot \frac{1}{2}}{P(X)}.$$

The Bayes procedure is to make the decision

$$\delta = \operatorname{sign}(P(S = 1|X) - P(S = -1|X)),$$

i.e. to guess that the state of the world is whichever state has higher posterior probability conditional on the votes X. So our goal is to show that $\delta = \operatorname{sign}(\sum_i X_i)$. To do this, note that $\delta = 1$ if and only if $\frac{P(S=1|X)}{P(S=-1|X)} > 1$. We can compute

$$\begin{aligned} \frac{P(S = 1|X)}{P(S = -1|X)} &= \frac{P(X|S = 1)}{P(X|S = -1)} = \frac{p^{\sum_i 1(X_i=1)}(1-p)^{\sum_i 1(X_i=-1)}}{p^{\sum_i 1(X_i=-1)}(1-p)^{\sum_i 1(X_i=1)}} \\ &= \left(\frac{p}{1-p}\right)^{\sum_i 1(X_i=1)} \left(\frac{1-p}{p}\right)^{\sum_i 1(X_i=-1)} \\ &= \left(\frac{p}{1-p}\right)^{\sum_i 1(X_i=1) - \sum_i 1(X_i=-1)} = \left(\frac{p}{1-p}\right)^{\sum_i X_i}. \end{aligned}$$

Since $\frac{p}{1-p} > 1$, the right side above will be greater than 1 if and only if $\sum_i X_i > 0$, and hence the Bayes procedure and the majority vote agree. □

References

[1] P. Bickel and K. Docksum. *Mathematical Statistics: Basic Ideas and Selected Topics.* Pearson Prentice Hall, 2nd edition, 2007.

[2] P. Billingsley. *Probability and Measure.* Wiley, 2nd edition, 1986.

[3] A. Dembo and O. Zeitouni. *Large deviations techniques and applications.* Springer-Verlag, 2nd edition, 2010.

[4] D.M. Estlund. Opinion Leaders, Independence, and Condorcet's Jury Theorem. *Theorey and Decision*, 36(2):131–162, 1994.

[5] O. Kallenberg. *Foundations of Modern Probability.* Springer, 2nd edition, 2002.