Some Remarks on the Probability of Cycles

Christian List

published as

"Appendix 3: An Implication of the *k*-option Condorcet jury mechanism for the probability of cycles"

in List and Goodin (2001)

Abstract. Standard results suggest that the probability of cycles should increase as the number of options increases and also as the number of individuals increases. These results are, however, premised on a so-called "impartial culture" assumption: any logically possible preference ordering is assumed to be as likely to be held by an individual as any other. The present chapter shows, in the three-option case, that given suitably systematic, however slight, deviations from an impartial culture situation, the probability of a cycle converges either to zero (more typically) or to one (less typically) as the number of individuals increases.

Although the pairwise Condorcet winner criterion may seem an attractive democratic decision procedure, it is famously threatened by Condorcet's paradox: pairwise majority voting may lead to cyclical collective preferences. But how probable is the occurrence of cycles?

An important body of literature addressing this question uses the so-called 'impartial culture' assumption (Gehrlein 1983). Given the $k!$ logically possible strict preference orderings, $P_1, P_2, ..., P_{k!}$, over k options, $x_1, x_2, ..., x_k$, it is assumed that all of these orderings are equally likely to be submitted by an individual, i.e. each individual has independent probabilities $p^*_1 = p^*_2 = ... = p^*_{k!} = 1/k!$ of submitting $P_1, P_2, ..., P_{k!}$ as his/her preference ordering, respectively. Given this perfect equiprobability assumption, the probability of the existence of a Condorcet winner *decreases* with increases in the number of individuals as well as with increases in the number of options. The larger the number of individuals, the harder it would seem to generate a Condorcet winning outcome.

This theoretical result is strikingly at odds with our empirical observations. Cycles are much less common in the real world than some of the social-choice-theoretic literature would lead us to expect (see Mackie 2000, for a critique of some famous purported empirical examples of cycles). But the result also seems hard to reconcile with another theoretical result. In a paper generalizing the Condorcet jury theorem from the case of majority voting over two options to the case of plurality voting over multiple options, List and Goodin (2001) have argued that, if there is a fact as to what the 'correct' best option is and each individual is more likely, however slightly, to track that fact in their preferences/votes, then several plausible social choice procedures (including the pairwise Condorcet winner criterion, the Borda count, and even plurality voting) converge in producing the 'correct' option as their unique winning outcome with a probality *increasing in the number of individuals*. In particular, under these assumptions the 'correct' option is also increasingly likely to emerge as the unique pairwise Condorcet winner as the number of individuals increases.

The response to this apparent clash of theoretical results is that the assumptions underlying the latter result break the 'impartial culture' assumption. Crucially, the assumption that each individual is more likely, however slightly, to choose the 'correct' option than any other is a violation of the assumption that all logically possible preference orderings are equally probable to occur. This raises the question of how much deviation from this equiprobability assumption is necessary to avoid the standard result on the probability of cycles. As we see in the present chapter, the mathematical mechanism that underlies the *k*-option Condorcet jury theorem (List and Goodin 2001) has an implication for this question.

Using the three-option case as a simple illustration, I now show that the impartial culture assumption can be seen as an extreme limiting case the slightest systematic deviation from which is already sufficient to circumvent the standard cycling result, provided the number of individuals is sufficiently large (a more technical and comprehensive treatment of some related results can be found in a very recent paper by Tangian 2000).

Suppose there are n individuals (n odd) and three options, x, y, z. For simplicity, we will only consider strict preference orderings. There are 6 logically possible such orderings of the options:

label	P_{X1}	P_{Y2}	P_{Z1}	P_{X2}	P_{Y1}	P_{Z2}
1st	z	z	y	y	x	x
2nd x	y	z	x	y	z	
3rd	y	x	x	z	z	y

Let $n(P_{X1})$, $n(P_{X2})$, $n(P_{Y1})$, $n(P_{Y2})$, $n(P_{Z1})$, $n(P_{Z2})$ be the numbers of individuals submitting orderings P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}, respectively. The vector $<n(P_{X1}), n(P_{X2}), n(P_{Y1}), n(P_{Y2}), n(P_{Z1}), n(P_{Z2})>$ is called an *anonymous preference profile*.

Proposition 1. (Nicholas Miller) The anonymous profile <$n(P_{X1})$, $n(P_{X2})$, $n(P_{Y1})$, $n(P_{Y2})$, $n(P_{Z1})$, $n(P_{Z2})$> generates a cycle under pairwise majority voting if and only if

$$\begin{aligned}&[\ [n(P_{X1}) > n(P_{X2})\ \&\ n(P_{Y1}) > n(P_{Y2})\ \&\ n(P_{Z1}) > n(P_{Z2})]\\ &\quad \textit{or}\ [n(P_{X1}) < n(P_{X2})\ \&\ n(P_{Y1}) < n(P_{Y2})\ \&\ n(P_{Z1}) < n(P_{Z2})]\ ]\\ &\&\ |n(P_{X1}) - n(P_{X2})| < n'/2\\ &\&\ |n(P_{Y1}) - n(P_{Y2})| < n'/2\\ &\&\ |n(P_{Z1}) - n(P_{Z2})| < n'/2,\end{aligned}$$

where $n' = |n(P_{X1}) - n(P_{X2})| + |n(P_{Y1}) - n(P_{Y2})| + |n(P_{Z1}) - n(P_{Z2})|$.

Now let $p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2}$ be the probabilities that an individual submits the orderings P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}, respectively (where the sum of the probabilities is 1). An *impartial culture* is the situation in which $p_{X1} = p_{X2} = p_{Y1} = p_{Y2} = p_{Z1} = p_{Z2}$.

Let X_{X1}, X_{X2}, X_{Y1}, X_{Y2}, X_{Z1}, X_{Z2} be the random variables whose values are the numbers of individuals with orderings P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}, respectively.

The joint distribution of X_{X1}, X_{X2}, X_{Y1}, X_{Y2}, X_{Z1}, X_{Z2} is a multinomial distribution with the following probability function:

$$P(X_{X1}=n_{X1}, X_{X2}=n_{X2}, \ldots, X_{Z2}=n_{Z2}) = \frac{n!}{n_{X1}!\ n_{X2}!\ \ldots\ n_{Z2}!}\, p_{X1}^{\ n_{X1}}\, p_{X2}^{\ n_{X2}} \ldots p_{Z2}^{\ n_{Z2}}.$$

Proposition 2. Suppose

$$\begin{aligned}&[\ [p_{X1} < p_{X2}\ \textit{or}\ p_{Y1} < p_{Y2}\ \textit{or}\ p_{Z1} < p_{Z2}]\\ &\quad \&\ [p_{X1} > p_{X2}\ \textit{or}\ p_{Y1} > p_{Y2}\ \textit{or}\ p_{Z1} > p_{Z2}]\ ]\\ &\textit{or}\ |p_{X1} - p_{X2}| > n'/2\\ &\textit{or}\ |p_{Y1} - p_{Y2}| > n'/2\\ &\textit{or}\ |p_{Z1} - p_{Z2}| > n'/2,\end{aligned}$$

where $n' = |p_{X1} - p_{X2}| + |p_{Y1} - p_{Y2}| + |p_{Z1} - p_{Z2}|$. Then the probability that there will be no cycle under pairwise majority voting tends to 1 as n tends to infinity.

Sketch proof. Consider the vector of random variables $\underline{X}^* = \langle X^*_{X1}, X^*_{X2}, X^*_{Y1}, X^*_{Y2}, X^*_{Z1}, X^*_{Z2}\rangle$, where, for each $i \in \{X1, X2, Y1, Y2, Z1, Z2\}$, $X^*_i := X_i/n$. The joint distribution of the X^*_i is a multinomial distribution with mean vector $\underline{p} = \langle p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2}\rangle$ and with variance-covariance matrix $\Sigma = (s_{ij})$, where, for each $i, j \in \{X1, X2, Y1, Y2, Z1, Z2\}$, $s_{ij} = p_i(1-p_i)$ if $i=j$ and $s_{ij} = -p_i p_j$ if $i \neq j$. Again by the central limit theorem, for large n, $(\underline{X}^*-\underline{p})\sqrt{(n)}$ has an approximate multivariate normal distribution $N(\underline{0}, \Sigma)$, and $\underline{X}^*-\underline{p}$ has an approximate multivariate normal distribution $N(\underline{0}, {}^1/_n\, \Sigma)$. Let $f_n : \mathbf{R}^6 \rightarrow \mathbf{R}$ be the corresponding density function for $\underline{X}^*-\underline{p}$. From proposition 1 we can infer, using this density function, that the probability that there will be no cycle under majority voting is given by $\int_{\underline{t}\in W} f_n(\underline{t})d\underline{t}$, where

$W := \{\underline{t} = \langle t_{X1}, t_{X2}, t_{Y1}, t_{Y2}, t_{Z1}, t_{Z2}\rangle \in \mathbf{R}^6 :$

$[\,[\,p_{X1}+t_{X1} < p_{X2}+t_{X2}$ *or* $p_{Y1}+t_{Y1} < p_{Y2}+t_{Y2}$ *or* $p_{Z1}+t_{Z1} < p_{Z2}+t_{Z2}]$

$\&\ [p_{X1}+t_{X1} > p_{X2}+t_{X2}$ *or* $p_{Y1}+t_{Y1} > p_{Y2}+t_{Y2}$ *or* $p_{Z1}+t_{Z1} > p_{Z2}+t_{Z2}]\,]$

or $|(p_{X1}+t_{X1}) - (p_{X2}+t_{X2})| > n'/2$

or $|(p_{Y1}+t_{Y1}) - (p_{Y2}+t_{Y2})| > n'/2$

or $|(p_{Z1}+t_{Z1}) - (p_{Z2}+t_{Z2})| > n'/2$,

where $n' = |(p_{X1}+t_{X1}) - (p_{X2}+t_{X2})| + |(p_{Y1}+t_{Y1}) - (p_{Y2}+t_{Y2})| + |p_{Z1}+t_{Z1}) - (p_{Z2}+t_{Z2})|\}$.

Note that, by assumption, $\underline{0} \in W$, and since all relevant inequalities satisfied by $p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2}$ are strict, there exists an $\varepsilon > 0$ such that $S_{\underline{0},\varepsilon} \subseteq W$, where $S_{\underline{0},\varepsilon}$ is a sphere around $\underline{0}$ with radius ε. Then, since f_n is nonnegative, $\int_{\underline{t}\in W} f_n(\underline{t})d\underline{t} \geq \int_{\underline{t}\in S_{\underline{0},\varepsilon}} f_n(\underline{t})d\underline{t}$. But, as f_n is the density function corresponding to $N(\underline{0}, {}^1/_n\, \Sigma)$, $\int_{\underline{t}\in S_{\underline{0},\varepsilon}} f_n(\underline{t})d\underline{t} \rightarrow 1$ as $n \rightarrow \infty$, and hence $\int_{\underline{t}\in W} f_n(\underline{t})d\underline{t} \rightarrow 1$ as $n \rightarrow \infty$, as required. **Q.E.D.**

Note that the condition of proposition 2 is already satisfied if at least one of $p_{X1} < p_{X2}$, $p_{Y1} < p_{Y2}$, $p_{Z1} < p_{Z2}$ and at least one of $p_{X1} > p_{X2}$, $p_{Y1} > p_{Y2}$, $p_{Z1} > p_{Z2}$ are satisfied. For instance, the condition is satisfied if $p_{X1} = 1/6 - \varepsilon$, $p_{Y1} = 1/6 + \varepsilon$ and $p_{X2} = p_{Y2} = p_{Z1} = p_{Z2} = 1/6$.

Proposition 2 implies that, given suitable systematic, however slight, deviations from an impartial culture, the probability that there will be a cycle under pairwise majority voting vanishes as the number of individuals increases.

The mechanism underlying this result is formally similar to the mechanism underlying the k-option Condorcet jury theorem. If $p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2}$ are the probabilities that an individual submits the orderings P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}, respectively, then np_{X1}, np_{X2}, np_{Y1}, np_{Y2}, np_{Z1},

np_{Z2} are the expected frequencies of these orderings amongst the n orderings submitted by an electorate of n individuals. If n is small, the actual frequencies may differ substantially from this pattern, but, as n increases, the actual frequencies approximate the expected frequencies increasingly closely in relative terms, by the law of large numbers. In particular, provided the probabilities satisfy the condition of proposition 2, the actual anonymous profile $<n(P_{X1}), n(P_{X2}), n(P_{Y1}), n(P_{Y2}), n(P_{Z1}), n(P_{Z2})>$ is thus increasingly likely to satisfy the negation of the condition of proposition 1 and hence decreasingly likely to generate a cycle.

However, if the probabilities deviate systematically from an impartial culture so as to replicate the pattern of Condorcet's paradox, the probability that there will be a cycle under pairwise majority voting tends to 1 as n tends to infinity.

Proposition 3. Suppose

$$
\begin{aligned}
&[\ [p_{X1} > p_{X2} \ \&\ p_{Y1} > p_{Y2} \ \&\ p_{Z1} > p_{Z2}] \\
&\quad \textit{or}\ [p_{X1} < p_{X2} \ \&\ p_{Y1} < p_{Y2} \ \&\ p_{Z1} < p_{Z2}]\ ] \\
\&\ & |p_{X1} - p_{X2}| < n'/2 \\
\&\ & |p_{Y1} - p_{Y2}| < n'/2 \\
\&\ & |p_{Z1} - p_{Z2}| < n'/2,
\end{aligned}
$$

where $n' = |p_{X1} - p_{X2}| + |p_{Y1} - p_{Y2}| + |p_{Z1} - p_{Z2}|$. Then the probability that there will be a cycle under pairwise majority voting tends to 1 as n tends to infinity.

Proof analogous to the proof of proposition 2.

An impartial culture is a rather unstable limiting case. We have seen that in any ε-neighbourhood of an impartial culture there are situations, as described by proposition 2, in which the probability of the occurrence of a cycle converges to 0 as the number of individuals increases. Likewise, there are situations, as described by proposition 3, in which the probability of the occurrence of a cycle converges to 1 as the number of individuals increases. It is an empirical question which of these situations is the more common one. Logically, the 'mostly disjunctive' condition of proposition 2 is less demanding than the 'mostly conjunctive' condition of proposition 3. Moreover, given the lack of empirical evidence of cycles, the condition of proposition 2 is arguably the empirically more plausible one.

Bibliography

Gehrlein, W. (1983) "Condorcet’s Paradox," *Theory and Decision* 15, 161-197.

Mackie, Gerry (2000) *The Coherence of Democracy*, unpublished manuscript, St. John's College, Oxford.

List, C., and Goodin, R. E. (2001) "Epistemic Democracy: Generalizing the Condorcet Jury Theorem," *Journal of Political Philosophy* 9(3), pp. 277-306.

Tangian, A. S. (2000) "Unlikelihood of Condorcet's paradox in a large society," *Social Choice and Welfare* 17, 337-365.