Soc Choice Welfare (1998) 15: 481–488

Social Choice and Welfare

When is Condorcet's Jury Theorem valid?

Daniel Berend[1], **Jacob Paroush**[2]

[1] Department of Mathematics and Computer Science, Ben-Gurion University, Beer-Sheva 84105, Israel
[2] Department of Economics, Bar-Ilan University, Ramat-Gan 52900, Israel

Received: 23 January 1997/Accepted: 8 March 1997

Abstract. Existing proofs of Condorcet's Jury Theorem formulate only sufficient conditions for its validity. This paper provides necessary and sufficient conditions for Condorcet's Jury Theorem. The framework of the analysis is the case of heterogeneous decisional competence, but the independence assumption is maintained.

1. Introduction

After the discovery of Condorcet's writings by Black (1958), the Condorcet's Essay has been recognized and appreciated as an important origin of social choice (see Urken 1991).

Being an enthusiastic supporter of the democratic regime, Condorcet believed that a group of individuals facing a binary choice and utilizing a simple majority rule would be likely to make the correct choice. Moreover, this likelihood would tend to complete certainty as the number of members of the group tends to infinity (see Baker 1976). A Condorcet's Jury Theorem (hereafter CJT) is a formulation of a sufficient condition (or conditions) that substantiates this belief. The simplest version of CJT suggests the condition that each member of the group has a competence $p > \frac{1}{2}$ to decide correctly and individuals vote independently, in the statistical sense. For a discussion of CJT see Miller (1986), Grofman and Feld (1988) and Young (1988, 1995).

Recently, there have been several attempts to generalize the popular version of CJT. Berg (1993a,b) and Ladha (1992, 1995) relax the independence assumption and allow for correlated votes. Grofman et al. (1983), Owen et al. (1989) and Paroush (1998) consider distribution-free team members' competence levels. Austen-Smith et al. (1996) analyze the case of insincere voting, and Louis et al. (1996) extend the dichotomous setup to a polychotomous one. The common denominator of all these studies is that they formulate only sufficient conditions for the above Condorcet's belief. In contrast, this paper presents a CJT with necessary and sufficient conditions. We adopt the dichotomous choice model with independent and sincere voting, but without any restrictions on the distribution of the decisional competence of the team members. Within this framework we prove that (in all practical cases) $\lim_{n\to\infty}(\bar{p}_n - \frac{1}{2})\sqrt{n} = \infty$ is a necessary and sufficient condition for $\lim_{n\to\infty}\pi_n = 1$, where $\bar{p}_n$ is the arithmetic mean of the team members' decisional abilities and π_n is the likelihood that the entire team (with n members) would reach the correct decision while utilizing a simple majority rule. This result is significant especially on the background of the example presented in Paroush (1997) showing that we do not necessarily have $\lim_{n\to\infty}\pi_n = 1$ even if $p_n > \frac{1}{2}$ for all n.

Section 2 presents the generalized CJT, Sect. 3 – the proof (along with a few more results), and in Sect. 4 we provide some examples.

2. The generalized CJT

Consider a team of n voters (jurors, decision makers) facing a binary choice. One of the alternatives is assumed to be objectively correct, but the team's members may have different abilities (competences) to identify this alternative. Denote the competence of the ith member of the team, namely his probability to decide correctly, by p_i, $i = 1, 2, \dots, n$. Throughout the paper we assume the voters to be independent. Intuitively, if the p_i's are "significantly" larger than $\frac{1}{2}$, a decision rule based on the simple majority rule will "most likely" lead to the right choice. Moreover, one would like to arrive at the same conclusion when assuming only that the average correctness probability

$$\bar{p}_n = \frac{1}{n}\sum_{i=1}^{n} p_i$$

is significantly larger than $\frac{1}{2}$. In the following we shall make these intuitive statements more precise. Denote by $P(n)$ the probability of making the right decision using the majority rule. More precisely, as the simple majority rule is defined only for an odd number n of individuals, we define $P(n)$ as the probability that (strictly) more than $\frac{n}{2}$ of the individuals will advocate the right decision.

Our main concern here is clarifying under what conditions the probability $P(n)$ becomes close to 1 as n becomes larger and larger. More formally, we assume that we have an infinite sequence $(p_i)_{i=1}^{\infty}$ of probabilities, representing an infinite collection of decision makers. The question is whether the probability of making correct decisions when using the majority rule, taking into account the opinions of more and more of those decision makers, converges to 1:

$$P(n) \underset{n\to\infty}{\longrightarrow} 1. \quad (1)$$

Obviously, the validity of (1) amounts to CJT. As mentioned already, (1) is known in the homogeneous case, where all the p_i's are identical, if $p_i = p > \frac{1}{2}$. Moreover, the same is basically known also in the heterogeneous case, where the p_i's are not necessarily identical, as long as $\bar{p}_n$ converges to a number strictly greater than $\frac{1}{2}$ (cf. Boland et al. 1989; Boland 1989; Owen et al. 1989). The main contribution of this paper is showing that (1) may hold even in situations where $p_n \underset{n\to\infty}{\longrightarrow} \frac{1}{2}$ (and in particular $\bar{p}_n \underset{n\to\infty}{\longrightarrow} \frac{1}{2}$). On the other hand, we show that the condition $\bar{p}_n > \frac{1}{2}$, or even $p_n > \frac{1}{2}$ for each n, is not sufficient to imply (1).

The formulation of the most general CJT, being cumbersome, will be postponed to the next section. Here we shall state it under an additional assumption, which actually holds in any practical situation. To this end, let us introduce the following definition. A sequence $(p_i)_{i=1}^{\infty}$ of probabilities is *reasonably balanced* if for some $\delta, \varepsilon > 0$ the inequality

$$\#(\{1 \leq i \leq n : \delta < p_i < 1 - \delta\}) > \varepsilon n \tag{2}$$

holds for all sufficiently large n (where $\#(S)$ denotes the cardinality of a finite set S). For the correctness probabilities of the team members to satisfy this condition means, roughly speaking, that some positive proportion of them consists of people who are neither "extremely smart" ($p_i \approx 1$) nor "extremely stupid" ($p_i \approx 0$).

Theorem 1. *If the sequence* (p_i) *is reasonably balanced, then* (1) *is valid if and only if*

$$\frac{\sum_{i=1}^{n} p_i - \frac{n}{2}}{\sqrt{n}} \underset{n\to\infty}{\longrightarrow} \infty.$$

Evidently, it is hard to imagine an unreasonably balanced sequence of probabilities in practice, so that Theorem 1 may be viewed as giving the "real" condition for CJT.

In fact, Theorem 1 is a corollary of an even more general CJT (Theorem 2 in the sequel), to be stated and proved in the next section.

3. An even more general CJT

We first formulate and prove the most general CJT for independent voters.

Theorem 2. (1) *is valid if and only if at least one of the following two conditions holds*:

1)
$$\frac{\sum_{i=1}^{n} p_i - \frac{n}{2}}{\sqrt{\sum_{i=1}^{n} p_i q_i}} \underset{n\to\infty}{\longrightarrow} \infty, \tag{3}$$

where $q_i = 1 - p_i$.

2) *For every sufficiently large* n

$$\#(\{i:\ 1 \le i \le n, p_i = 1\}) > \frac{n}{2}. \tag{4}$$

Proof. In both parts of the proof we shall utilize the sequence $(X_i)_{i=1}^{\infty}$ of random variables defined by

$$X_i = \begin{cases} 1, & \text{the } i\text{th individual chooses correctly,} \\ 0, & \text{otherwise.} \end{cases}$$

Then $X = \sum_{i=1}^{n} X_i$ is the number of individuals voting correctly, and:

$$E(X) = \sum_{i=1}^{n} p_i, \quad V(X) = \sum_{i=1}^{n} p_i q_i.$$

We start the proof with the sufficiency part. Obviously, the second condition of the theorem implies (1). To prove the sufficiency of the first condition we note that from Chebychev's Inequality it follows readily that

$$\begin{aligned} 1 - P(n) = P\Big(X < \frac{n}{2}\Big) &\le P\left(\left|X - \sum_{i=1}^{n} p_i\right| \ge \sum_{i=1}^{n} p_i - \frac{n}{2}\right) \\ &\le \frac{V(X)}{\left(\sum_{i=1}^{n} p_i - \frac{n}{2}\right)^2} = \frac{\sum_{i=1}^{n} p_i q_i}{\left(\sum_{i=1}^{n} p_i - \frac{n}{2}\right)^2} \underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

and consequently

$$P(n) \underset{n\to\infty}{\longrightarrow} 1.$$

To prove the necessity of the conditions, we distinguish between two cases. Suppose first that

$$\sum_{i=1}^{n} p_i q_i = \infty. \tag{5}$$

We claim that (3) is satisfied. In fact, from (5) it follows that the sequence (X_i) satisfies Lindeberg's condition, and therefore the central limit theorem. (See, for instance, Feller (1971), Theorem VIII.4.3.) Suppose (3) does not hold. Then, for a suitable constant C, we have

$$\frac{\sum_{i=1}^{n} p_i - \frac{n}{2}}{\sqrt{\sum_{i=1}^{n} p_i q_i}} < C$$

for infinitely many integers n. Hence, denoting by Φ the normal distribution function, we obtain for an arbitrary fixed $\varepsilon > 0$ and sufficiently large such n:

$$\begin{aligned} P(n) = P\left(\sum_{i=1}^{n} X_i > \frac{n}{2}\right) &= P\left(\sum_{i=1}^{n} (X_i - p_i) > \frac{n}{2} - \sum_{i=1}^{n} p_i\right) \\ &= P\left(\frac{\sum_{i=1}^{n} (X_i - p_i)}{\sqrt{\sum_{i=1}^{n} p_i q_i}} > \frac{\frac{n}{2} - \sum_{i=1}^{n} p_i}{\sqrt{\sum_{i=1}^{n} p_i q_i}}\right) \le 1 - \Phi(-C) + \varepsilon. \end{aligned}$$

The right hand side may be made less than 1, which contradicts (1).

We may assume consequently that

$$\sum_{i=1}^{\infty} p_i q_i = V < \infty.$$

It will be convenient to split this case into two subcases, depending on the cardinality of the set

$$E = \left\{ i : \frac{1}{2} \le p_i < 1 \right\}.$$

If E is finite, define a sequence $(\eta_i)_{i=1}^{\infty}$ by:

$$\eta_i = \begin{cases} 1, & p_i = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then:

$$P(\{X_i = \eta_i,\ i = 1, 2, \ldots\}) = \prod_{i \in E} q_i \cdot \prod_{p_i < 1/2} q_i.$$

Now the first term on the right hand side is non-zero due to the finiteness of E, while the second is non-zero since

$$\sum_{p_i < 1/2} p_i \le \sum_{p_i < 1/2} p_i \cdot 2q_i \le 2V < \infty.$$

Consequently, if (4) is not satisfied for a certain n, then

$$P(n) \le 1 - P(\{X_i = \eta_i,\ i = 1, 2, \ldots, n\}) \le 1 - P(\{X_i = \eta_i,\ i = 1, 2, \ldots\}) < 1,$$

and therefore (4) must hold from some place on.

It remains to deal with the case where E is infinite. Suppose (3) is not satisfied. Then for infinitely many integers n we have

$$\sum_{i=1}^{n} p_i - \frac{n}{2} < C, \tag{6}$$

where C is an appropriate constant. Take $i_1, i_2, \ldots, i_r \in E$ with $r > 2\left(\sqrt{2V} + C\right)$. Denote $E' = \{i_1, i_2, \ldots, i_r\}$. If $n > \max_{1 \le j \le r} i_j$ satisfies (6) then:

$$\begin{aligned} 1 - P(n) &= P\left(\sum_{i=1}^{n} X_i \le \frac{n}{2}\right) \ge P\left(X_{i_1} = X_{i_2} = \cdots = X_{i_r} = 0,\ \sum_{i=1}^{n} X_i \le \frac{n}{2}\right) \\ &= \prod_{j=1}^{r} q_{i_j} P\left(\sum_{1 \le i \le n, i \notin E'} X_i \le \frac{n}{2}\right). \end{aligned} \tag{7}$$

Now:

$$\begin{aligned} P\left(\sum_{1\le i\le n, i\notin E'} X_i \le \frac{n}{2}\right) &= P\left(\sum_{1\le i\le n, i\notin E'} (X_i - p_i) \le \frac{n}{2} - \sum_{1\le i\le n, i\notin E'} p_i\right) \\ &\ge P\left(\left|\sum_{1\le i\le n, i\notin E'} (X_i - p_i)\right| \le \frac{n}{2} - \sum_{1\le i\le n, i\notin E'} p_i\right) \\ &\ge 1 - \frac{\sum_{1\le i\le n, i\notin E'} p_i q_i}{\left(\frac{n}{2} - \sum_{1\le i\le n, i\notin E'} p_i\right)^2} \\ &\ge 1 - \frac{V}{2V} = \frac{1}{2}. \end{aligned} \tag{8}$$

Clearly, (7) and (8) are incompatible with (1), which concludes the proof of the last case. □

As we shall see subsequently (Example 1 in the next section), condition 1 in Theorem 2 is not necessary for CJT to hold in general. However, it is necessary under "most" circumstances. In fact, going carefully over the proof of Theorem 2, one finds that the following is true.

Theorem 3. *If infinitely many of the probabilities* $(p_i)_{i=1}^{\infty}$ *belong to the interval* $[\frac{1}{2}, 1)$, *then condition* 1 *in Theorem* 2 is necessary and sufficient for (1) *to be valid.*

Now we can conclude the proof of Theorem 1. In fact, on the one hand we always have

$$\sum_{i=1}^{n} p_i q_i \le \frac{n}{4},$$

so that the sufficiency part of Theorem 1 follows from Theorem 2. On the other hand, since (p_i) is reasonably balanced, we obtain

$$\sum_{i=1}^{n} p_i q_i \ge \delta\sqrt{\varepsilon(1-\varepsilon)}\sqrt{n}$$

(where $\delta, \varepsilon > 0$ are as in (2)). Theorem 2 and its proof now yield easily the necessity part of Theorem 1 as well.

4. Examples

An immediate consequence of Theorem 2 is the following main result of Paroush (1998).

Corollary. *If* $p_i \ge \frac{1}{2} + \varepsilon$ *for each* i, *where* $\varepsilon > 0$ *is fixed, then* (1) *is valid.*

In fact, we have

$$\frac{\sum_{i=1}^{n} p_i - \frac{n}{2}}{\sqrt{\sum_{i=1}^{n} p_i q_i}} \geq \frac{n\varepsilon}{\sqrt{n\left(\frac{1}{2}+\varepsilon\right)\left(\frac{1}{2}-\varepsilon\right)}} = \frac{\varepsilon}{\sqrt{\left(\frac{1}{2}+\varepsilon\right)\left(\frac{1}{2}-\varepsilon\right)}} \sqrt{n} \underset{n\to\infty}{\longrightarrow} \infty,$$

which implies the corollary.

However, the probabilities may be quite closer to $\frac{1}{2}$ than required in the corollary, and even converge to $\frac{1}{2}$ with the same conclusion still holding.

Example 1. Let $p_i = \frac{1}{2} + \frac{1}{i^\theta}$ for sufficiently large i. If $\theta < \frac{1}{2}$ then for an appropriate C and sufficiently large n

$$\frac{\sum_{i=1}^{n} p_i - \frac{n}{2}}{\sqrt{\sum_{i=1}^{n} p_i q_i}} > \frac{\sum_{i=1}^{n} \frac{1}{i^\theta} + C}{\sqrt{\frac{n}{4}}} > n^{\frac{1}{2}-\theta} \underset{n\to\infty}{\longrightarrow} \infty,$$

so that (3) holds.

The following example shows that (3) is not necessary for (1) to hold.

Example 2. Suppose $p_1 \neq 0, 1$ is arbitrary, $p_2 = p_3 = p_4 = 1$, $p_{2i-1} = 0$ for $i \geq 3$ and $p_{2i} = 1$ for $i \geq 3$. Then, with probability 1, the majority rule will lead to the correct decision for every $n \geq 3$. However, the expression $\frac{\sum_{i=1}^{n} p_i - \frac{n}{2}}{\sqrt{\sum_{i=1}^{n} p_i q_i}}$ assumes only two values as n varies, and in particular does not diverge to ∞.

References

Austen-Smith D, Banks J (1996) Information, aggregation, rationality and CJT. Amer Polit Sci Rev 90(1): 34–45

Baker KM (Ed) (1976) Condorcet: Selected writings. The Bobbs-Mervill, Indiapolis

Black D (1958) The Theory of Communities and Elections. Cambridge University Press, Cambridge

Berg S (1993a) Condorcet's Jury Theorem, Dependence Among Jurers. Soc Choice Welfare 10: 87–95

Berg S (1993b) Condorcet's Jury Theorem revisited. Eur J Polit Econ 9: 437–446

Boland PJ (1989) Majority systems and the Condorcet Jury Theorem. Statistician 38: 181–189

Boland PJ, Proschan F, Tong YL (1989) Modelling dependence in simple and indirect majority systems. J Appl Probab 26: 81–88

Condorcet NC de (1785) Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris

Feller W (1971) An introduction to probability theory and its applications, Vol. 2, 2nd edn. John Wiley & Sons, New York, London

Grofman B, Feld SL (1988) Rousseau's general will: A Condorcet perspective. Amer Polit Sci Rev 82: 567–576

Grofman B, Owen G, Feld SL (1983) Thirteen theorems in search of the truth. Theory Decision 15: 261–278

Ladha KK (1992) The Condorcet Jury Theorem, free speech and correlated votes. Amer J Polit Sci 36: 617–634

Ladha KK (1995) Information polling through majority rule voting: Condorcet's Jury Theorem with correlated votes. J Econ Behavior Organizat 26: 353–372

Louis L, Ching YS (1996) Majority vote of even and odd experts in a polychotomous choice situation. Theory Decision 41: 13–36
Miller NR (1986) Information, electorates, and democracy: Some extensions and interpretations of the Condorcet Jury Theorem. In: Grofman B, Owen G (eds) Information pooling and group decision making. JAI Press, Greenwich, CT
Owen G, Grofman B, Feld SL (1989) Proving a distribution-free generalization of the Condorcet Jury Theorem. Math Soc Sci 17: 1–16
Paroush J (1998) Stay away from fair coins: A Condorcet's Jury Theorem. Soc Choice Welfare 15: 15–20
Young HP (1988) Condorcet's Theory of voting. Amer Polit Sci Rev 82: 1231–1244
Young HP (1995) Optimal voting rules. Econ Perspect 9: 51–64