FRANZ DIETRICH and CHRISTIAN LIST

A MODEL OF JURY DECISIONS WHERE ALL JURORS HAVE THE SAME EVIDENCE

24 September 2002

ABSTRACT. In the classical Condorcet jury model, different jurors' votes are independent random variables, where each juror has the same probability $p>1/2$ of voting for the correct alternative. The probability that the correct alternative will win under majority voting converges to 1 as the number of jurors increases. Hence the probability of an incorrect majority vote can be made arbitrarily small, a result that may seem unrealistic. A more realistic model is obtained by relaxing the assumption of independence and relating the vote of every juror to the same "body of evidence". In terms of Bayesian trees, the votes are direct descendants not of the true state of the world ('guilty' or 'not guilty'), but of the "body of evidence", which in turn is a direct descendant of the true state of the world. This permits the possibility of a misleading body of evidence. Our main theorem shows that the probability that the correct alternative will win under majority voting converges to the probability that the body of evidence is not misleading, which may be strictly less than 1.

KEY WORDS: Condorcet jury theorem, conditional independence, interpretation of evidence, Bayesian trees

1. INTRODUCTION

Suppose a jury has to reach a decision on whether or not a defendant is guilty. There are two possible states of the world: $x = 1$ (the defendant is guilty) and $x = 0$ (the defendant is not guilty). Given that the state of the world is x, each juror has the same probability (competence) $p > ½$ of voting for x and the votes of different jurors are independent from each other. Then the probability that a majority of jurors votes for x, given the state of the world x, converges to 1 as the number of jurors increases. This is the classical Condorcet jury theorem. The result implies that, so long as the number of jurors is sufficiently large, the reliability of a majority decision can be made arbitrarily close to certainty.

Something about this result may seem puzzling. What if, for instance, all jurors are tricked by the same evidence, which seems ever so compelling? What if, against all odds, the unbelievable happens and the wind blows a hair of an innocent person to the exact scene of the crime and the jurors all come to believe that there is no plausible way in which that hair could ever have arrived there without the person? What if the evidence is so sparse that, no matter how many jurors are consulted, the sheer lack of evidence makes it impossible to solve a case conclusively?

In response, it might be argued that all we need to do to rule out each of these scenarios is to increase the number of jurors sufficiently. Suppose each juror views the scene of the crime from a different perspective. Or suppose each juror has obtained a separate item of evidence about the crime, where all such items of evidence are perfectly independent from each other. And suppose further that we can find as many jurors (and items of evidence) as we like that all satisfy this description. Then the jury might indeed be able to reach a correct decision with a probability approaching 1. In that case the jury would be able to aggregate arbitrarily many independent items of evidence into a single overall verdict. Call this case A. But often there are not arbitrarily many *independent* items of evidence. Rather, the jury as a whole reviews the *same* body of evidence, and each juror has to decide whether he or she believes that this body of evidence establishes guilt beyond any reasonable doubt. Call this case B.

Case A does indeed satisfy the conditions of the classical Condorcet jury theorem. But case B does not, or so we will argue. We develop a new model for formalizing case B. In our model, different jurors are not independent *conditional on the state of the world* (as in the classical Condorcet jury model), but they are independent *only conditional on the evidence*. The model shows that, no matter how many jurors are consulted and how competent each juror is, the overall reliability of the jury is always bounded above by the probability that the evidence is not misleading. And that probability may differ from case to case and may in principle rule out the kind of reliability predicted by the classical Condorcet jury theorem. We prove that, as the number of jurors increases, the probability that a majority of jurors votes for x, given that the state of the world is x, *converges* to the probability that the evidence about this state of the world is not misleading, a value that is typically strictly less than one. The results imply that, to the extent that real world jury decisions are more similar to case B than to case A, the classical Condorcet jury theorem fails to apply to real world jury decisions.

In the large literature on the Condorcet jury theorem (see, amongst many others, Grofman, Owen and Feld 1983; Young 1988; Austen-Smith and Banks 1996; Berend and Paroush 1998; List and Goodin 2001), only a few papers address dependencies between the votes of different jurors. Among the discussions of dependencies, there has been a focus on the role of opinion leaders – i.e. jurors on whose views the views of other jurors are dependent – (Grofman, Owen and Feld 1983; Nitzan and Paroush 1984; Boland 1989; Boland, Proschan and Tong; Estlund 1994) and on the effect of free speech on keeping correlations low (Lahda 1992). However, existing models of dependencies have usually preserved the classical result that the probability of a majority for the correct alternative converges to 1 as the number of jurors increases, so long as the votes of different jurors are not too highly correlated. In particular, these models do not impose an upper bound on the totality of evidence available to the jury, and they usually suggest that the key difference between the classical framework and a framework with dependencies (unless these dependencies are too great) lies in a different (i.e. slower) *speed of convergence*, but *not* in a different *limit*. Our model, on the other hand, shows that dependencies resulting from the use of the same body of evidence by all jurors may lead to convergence of the reliability of the jury verdict to a different *limit*, namely to the probability that the evidence is not misleading.

2. THE MODEL

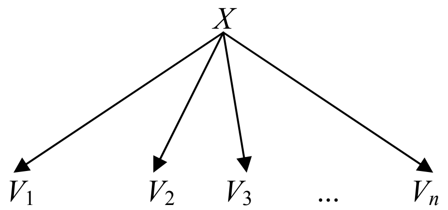
We assume there are n jurors, labelled $i = 1, 2, ..., n$. The state of the world is represented by a binary variable X which takes the value 1 for 'guilty' and 0 for 'not guilty'. The votes of the jurors are represented by the binary random variables V_1, V_2, ..., V_n, where each V_i takes the value 1 for a 'guilty' vote and 0 for a 'not guilty' vote. A juror makes a correct judgment if and only if the value of V_i coincides with the value of X. As a notational convention, we use capital letters to denote random variables and corresponding small letters to denote particular values.

The classical Condorcet jury theorem assumes the following:

Competence ($\mathbf{C_{classical}}$). For each $x\in\{0, 1\}$ and all jurors $i = 1, 2, ..., n$, $p{:=}P(V_i{=}x|X{=}x){>}½$.

Independence ($\mathbf{I_{classical}}$). For each $x\in\{0, 1\}$, V_1, V_2, ..., V_n are independent from each other, given the state of the world x.

DIAGRAM 1: BAYESIAN TREE FOR THE CLASSICAL CONDORCET JURY MODEL



The (conditional) independence assumption ($C_{classical}$) can be illustrated by the Bayesian tree shown in diagram 1. Each juror receives an independent signal about the state of the world and votes exactly on the basis of that signal. The signal is noisy, but it is biased towards the truth, in so far as $p > ½$. The classical Condorcet jury theorem states that majority voting over the independent signals received by different jurors reduces the noise.

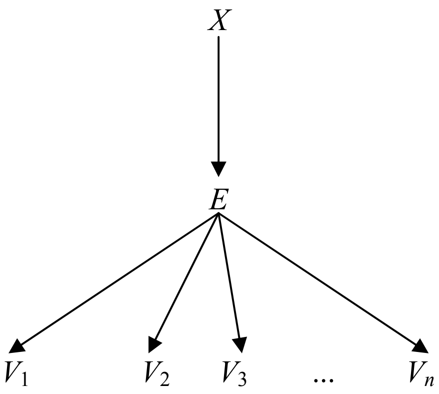
Let $V = \Sigma_{i=1,\ldots,n}V_i$. Then $V > n/2$ corresponds to a majority for 'guilty', and $V < n/2$ corresponds to a majority for 'not guilty'.

THEOREM 1. (Condorcet jury theorem) *If we have ($C_{classical}$) and ($I_{classical}$), then the probabilities* $P(V{>}n/2|X{=}1)$ *and* $P(V{<}n/2|X{=}0)$ *converge to* 1 *as n tends to infinity.*

The new model gives up the assumption that each juror receives an independent signal about the state of the world. Instead, the state of the world generates a single overall signal, E, interpreted as a body of evidence; E is a random variable which takes values in some set $\mathbf{E}$ of all possible bodies of evidence. The value of E can be interpreted as the totality of available information about the state of the world, including for instance the testimony of a witness or the particular appearance of the defendant in court (relaxed or stressed, smiling or serious etc.). What matters is not the particular nature of the signal, which will usually have a complex form, but the fact that every juror observes the *same* signal. The probability distribution of E depends on the state of the world. The distribution of E given guilt (x=1) is different from that given innocence (x=0). For instance, in the case of guilt we might expect a greater probability of observing stressed behaviour on the part of the defendant than in the case of innocence.

This model captures not only the *fact* that in real world jury decisions the available evidence is usually finite and limited, but also the common *legal norm* that jurors are not allowed to obtain or use any evidence other than the one presented in the actual courtroom, or to discuss the case with any persons other than the other jury members.

DIAGRAM 2: BAYESIAN TREE FOR THE NEW MODEL



The structure of the new model can be illustrated by the Bayesian tree in diagram 2. All jurors base their votes solely on the same value e of E. The difference in judgments between the jurors results not from different signals, but from different interpretations of the same signal e. One juror might interpret a smile on the face of the defendant as a sign of innocence, while another juror might give the opposite interpretation. We also allow the case that not all jurors have observed the entire signal e. For instance, some jurors might have missed the smile of the defendant. Hence, what matters is not that all jurors base their decision on the "full" signal e, but that the information used by each juror is *contained* in e. The signal e is thus interpreted as the *maximal* available information, i.e. the information pool out of which each juror's individual information is taken, whether the jury contains 5 or 100 jurors. In a slight abuse of language, a juror's interpretation of e is intended to capture both the fact that the juror might have received only part of the information contained in e and the fact that the juror has an individual way of reading that information.[1]

How "competent" is each juror? While in the classical Condorcet jury model competence was modelled by each juror's probability $p > 1/2$ of making a *correct decision*, *conditional on the state of the world*, we here propose to model competence in terms of the probability of giving an *ideal interpretation* of the evidence, *conditional on that evidence*. Specifically, we suppose that for any body of evidence e in $\mathbf{E}$ there exists an "ideal" interpretation or vote, $f(e)$, that would be given by a hypothetical ideal observer (interpreter) of e. This ideal observer does not know the true state of the world x, but gives the ideal (best possible) interpretation of the available body of evidence. Again, $f(e) = 1$ means that the ideal observer would vote for 'guilty' and $f(e) = 0$ means that the ideal observer would vote for 'not guilty'. We call $f(e)$ the "ideal" vote, by contrast to the "correct" vote, which is the vote matching the true state of the world.[2] While knowledge

[1] This model applies in great generality. It even allows us to capture the case of deliberation among the jurors, by interpreting the "body of evidence" e very broadly. In that case, e would include not only the "evidence" in a narrow sense, but also the entire jury deliberation process, up to the point where a vote is taken.

[2] To make the notion of an ideal vote tractable, we might give a Bayesian account, which we here illustrate for the special case where the set $\mathbf{E}$ of all possible bodies of evidence is countable. Suppose that, by knowing the evidence-generating stochastic process, the ideal observer knows the probabilities $P(E=e|x=1)$ and $P(E=e|x=0)$. Suppose, further, that the ideal observer assigns some (typically low) prior probability

of the true state x would allow a correct vote, the ideal vote results from the best possible interpretation of e. Crucially, the ideal vote and the correct vote will differ in the case of misleading evidence, for instance when the wind blows a hair of an innocent person to the exact scene of the crime (and when this unfortunate person has no other alibi etc.).

We can now formulate the assumptions of our modified jury model. Our competence and independence assumptions resemble the assumptions of the classical model, but with the important difference that we conditionalize on the evidence e rather than on the state of the world x (as in the classical model).

Common signal (S). For each $e \in \mathrm{E}$ and each $x \in \{0, 1\}$, the joint probability distribution of $V_1, V_2, \ldots, V_n$ given both $E=e$ and $X=x$ is the same as that given just $E=e$.

Informally, the jurors' votes depend on the true state of the world only *through* the available evidence. Once the evidence is given, what the true state of the world is makes no difference to the probability distribution of the jurors' votes.

Competence (C). For each $e \in \mathrm{E}$, each $x \in \{0, 1\}$ and all jurors $i = 1, 2, ..., n$, $p_e := P(V_i = f(e) | E=e) > ½$. The value of p_e depends on e.

Informally, the probability that juror i's vote matches the ideal vote $f(e)$ given the evidence e is a number $p_e > ½$ that is identical for all jurors.

Independence (I). For each $e \in \mathrm{E}$, $V_1, V_2, ..., V_n$ are independent from each other, given the evidence e.

Informally, once the evidence is given, the votes of different jurors are independent from each other.

In general, the competence p_e is a function of the evidence e. If the body of evidence e is easily interpretable, for instance in the case of overwhelming evidence for innocence, the probability that an individual juror's vote matches the ideal vote $f(e)=0$ might be high, say $p_e=0.95$, whereas if the body of evidence e is sparse or ambiguous that probability might be only $p_e=0.55$. Thus our notion of competence is in general a whole family $\{p_e : e \in \mathrm{E}\}$ of probabilities. The term 'competence' here corresponds to the ability to interpret the different possible bodies of evidence $e \in \mathrm{E}$ in a way that matches the ideal interpretation. For simplicity, one might prefer to replace (C) with the stronger (and less realistic) assumption of homogeneous competence, according to which p_e is identical for all possible $e \in \mathrm{E}$.

$r := P(x=1)$ to the proposition that the defendant is guilty. Then, using Bayes's theorem, the ideal observer is able to arrive at a posterior probability that the defendant is guilty, given the evidence e, i.e. $P(x=1|E=e) = rP(E=e|x=1) / (rP(E=e|x=1) + (1-r)P(E=e|x=0))$. Now the ideal observer might set a (normative) criterion of when he or she believes, beyond any reasonable doubt, that the defendant is guilty, given the evidence e. Specifically, the ideal observer might vote for 'guilty' (i.e. $f(e)=1$) if $P(x=1|E=e) > 1-\varepsilon$ (for some suitably chosen $\varepsilon > 0$) and for 'not guilty (i.e. $f(e)=0$) otherwise. Different values of the prior probability r correspond to what degree of belief the ideal observer assigns to the innocence of the defendant (namely $1-r$) *before* having seen any evidence; and different values of ε correspond to differentially demanding criteria of what degree of belief counts as 'beyond any reasonable doubt'. These considerations illustrate that the notion of an ideal interpretation is, in part, a normative notion.

Homogeneous competence (C*). For each $e \in \mathrm{E}$, each $x \in \{0, 1\}$ and all jurors $i = 1, 2, ..., n$, $p := P(V_i = f(e) | E = e) > ½$. The value of p does not depend on e.

3. THE PROBABILITY DISTRIBUTION OF THE JURY'S VOTE GIVEN THE TRUE STATE OF THE WORLD

We now derive the probability distribution of the jury's vote $V = \Sigma_{i=1,\ldots,n} V_i$ given the true state of the world (theorems 2 and 3). We see that, given guilt, the probability of a simple majority for guilt is at most as high as (and typically strictly smaller than) it is in the classical Condorcet model (corollary 1).

We first assume (S), (C*) and (I). Since (C*) holds, we further assume for simplicity that $\mathrm{E} = \{0, 1\}$, where, for each $e \in \mathrm{E}$, $f(e) = e$. Now, by (C*) and (I), given the evidence e, if e=1 then each juror's vote V_i has an independent Bernoulli distribution, with a probability p of $V_i = 1$ and a probability (1-p) of $V_i = 0$; if e=0 then each V_i also has an independent Bernoulli distribution, but with a probability p for $V_i = 0$ and a probability (1-p) for $V_i = 1$. Hence, given the evidence e, if e=1 the jury's vote $V = \Sigma_{i=1,\ldots,n} V_i$ has a Binomial distribution with parameters n and p, and if e=0 it has a Binomial distribution with parameters n and (1-p):

$$P(V=v|E=1) = \binom{n}{v} p^v (1-p)^{n-v}, \quad P(V=v|E=0) = \binom{n}{v} p^{n-v} (1-p)^v \qquad (*)$$

Now, the probability of obtaining precisely v out of n votes for 'guilty' given the true state of the world x is the following:

$$P(V=v|X=x) = P(V=v|E=1 \textit{ and } X=x)P(E=1|X=x) + P(V=v|E=0 \textit{ and } X=x)P(E=0|X=x).$$

By (S), conditionalizing on *both* $E=e$ *and* $X=x$ is equivalent to conditionalizing on $E=e$ only, so that:

$$P(V=v|X=x) = P(V=v|E=1)P(E=1|X=x) + P(V=v|E=0)P(E=0|X=x).$$

Since $f(E) = E$, (*) implies the following theorem:

THEOREM 2. *If we have (S), (C*) and (I), the probability of obtaining precisely v out of n votes for 'guilty' given the true state of the world x is*

$$P(V=v|X=x) = P(f(E)=1|X=x)\binom{n}{v} p^v (1-p)^{n-v} + P(f(E)=0|X=x)\binom{n}{v} p^{n-v} (1-p)^v.$$

Note that $P(f(E)=0|X=x) = 1 - P(f(E)=1|X=x)$.

Theorem 3 below implies that theorem 2 still holds if we relax the assumption that $\mathrm{E} = \{0, 1\}$ and consider a more general E.

If there is a non-zero probability of misleading evidence – specifically if $0 < P(f(E) \neq x | X=x) < 1$ – the jury's vote V given the state of the world x does not have a binomial distribution, in contrast to the classical Condorcet jury model. The reason for this is that the votes V_1, V_2, ..., V_n are not independent from each other given the state of the world x, but they are only independent from each other given the evidence e. The sum of *dependent* Bernoulli variables does not in general have a binomial distribution. If, on the other hand, the probability of misleading evidence is zero – i.e. $P(f(E)\neq x|X=x)=0$ – the probability in theorem 2 reduces to the one in the classical Condorcet jury model.

In the appendix we derive the probability $P(V=v|X=x)$ for the more general case where we assume (C) rather than (C*) and where the set of all possible bodies of evidence $\mathbf{E}$ can be more general. Note that since E is a random variable, E induces a random variable p_E which always takes as its value the competence p_e associated with the value e of E. To avoid confusion with the random variable E, we write the *expected value operator* as $Exp(.)$.

THEOREM 3. *If we have (S), (C) and (I), the probability of obtaining precisely v out of n votes for 'guilty' given the true state of the world x is*

$$P(V=v|X=x) = P(f(E)=1|X=x)\binom{n}{v}Exp(p_E^{\,v}(1-p_E)^{n-v}\,|f(E)=1 \textit{ and } X=x)$$

$$+ P(f(E)=0|X=x)\binom{n}{v}Exp(p_E^{\,n-v}(1-p_E)^{v}|f(E)=0 \textit{ and } X=x).$$

In theorems 2 and 3, by taking x=1 we get the probability of $V=v$ given guilt, and by taking x=0 we get the probability of $V=v$ given innocence. Also, by taking x=1 and summing the probabilities over all $v > n/2$, we obtain the probability of a simple majority for 'guilty' given guilt; and, by taking x=0 and summing the probabilities over all $v < n/2$, we obtain the probability of a simple majority for 'not guilty' given innocence.

If (C*) holds, we can deduce an interesting inequality. In the formula of theorem 2, assume that $v>n/2$ (more votes for 'guilty' than for 'not guilty'). Then

$$p^{n-v}(1-p)^{v} = p^{v}(1-p)^{n-v}((1-p)/p)^{2v-n} < p^{v}(1-p)^{n-v},$$

since $2v-n>0$ and $p>1/2$. Further since $P(f(E)=1|X=x)+P(f(E)=0|X=x)=1$, we deduce:

COROLLARY 1. *Suppose we have (S), (C*) and (I). Let* $v>n/2$. *Then the probability of obtaining precisely v out of n votes for 'guilty' given guilt satisfies*

$$P(V=v|X=1) \leq \binom{n}{v}p^{v}(1-p)^{n-v},$$

and the probability of obtaining a majority for 'guilty' given guilt satisfies

$$P(V>n/2|X=1) \leq \sum_{v>n/2}\binom{n}{v}p^{v}(1-p)^{n-v}.$$

Thus the probability of obtaining precisely v votes for 'guilty' given guilt in our model is less than or equal to that in Condorcet's model with the same competence parameter p. The same holds for the probability of obtaining a majority for 'guilty' given guilt. Similarly, one can show that in our model the probability of a majority for 'not guilty' given innocence is bounded above by the corresponding probability in the classical model. In both inequalities in corollary 1, equality holds if and only if the probability of misleading evidence is zero, i.e. if and only if $P(f(E)=0|X=1)=0$.

Informally, unless the evidence always "tells the truth", the jury in our model will reach a simple majority for the correct alternative with a lower probability than in the classical Condorcet jury model.

4. A MODIFIED CONDORCET JURY THEOREM

We now state our modified version of the Condorcet jury theorem (theorem 4). The first part of the theorem is concerned with the probability that the majority of jurors matches the ideal vote, and the second part with the probability that the majority of jurors matches the true state of the world.

THEOREM 4. *Suppose we have (S), (C) and (I).*

(i) *Let W be the number of jurors $i\in\{1, 2, ..., n\}$ such that $V_i = f(E)$. For each $x\in\{0,1\}$, $P(W > n/2|X=x)$ converges to 1 as n tends to infinity.*

(ii) *$P(V>n/2|X=1)$ converges to $P(f(E)=1|X=1)$ as n tends to infinity, and $P(V<n/2|X=0)$ converges to $P(f(E)=0|X=0)$ as n tends to infinity.*

Part (i) states that, given the state of the world, the probability that a simple majority of jurors matches the ideal interpretation of the evidence converges to 1 as n tends to infinity; part (ii) states that the probability that a simple majority of jurors matches the true state of the world converges to the probability that the ideal interpretation of the evidence is correct, i.e. that the evidence is not misleading. Theorem 4 immediately implies that, given the state of the world, the probability that there will be no simple majority for the *ideal* interpretation of the evidence converges to 0 as n tends to infinity; and the probability that there will be no simple majority that matches the true state of the world converges to the probability that the evidence is misleading, i.e. that the ideal interpretation of the evidence is incorrect.

This theorem allows the interpretation that, by increasing the size of the jury, it is possible to approximate the ideal interpretation of the evidence, no more and no less. The problem of insufficient or misleading evidence cannot be avoided by adding new jurors. Irrespective of the size of the jury, the probability of a correct majority decision remains bounded above by the probability that the evidence "tells the truth", i.e. that it leads to an ideal interpretation which matches the state of the world. Since there is typically a nonzero probability of misleading evidence – i.e. a nonzero probability that the evidence, even when ideally interpreted, points to 'guilt' when the defendant is innocent or vice-versa – there is also a nonzero probability that the jury will fail to track the truth, regardless of how large the jury is and what the competence parameters p_e are in assumption (C).[3]

[3] It is possible to prove a slightly stronger result than the one in theorem 4. Given the state of the world x, the ratio V/n converges with probability 1 to the random variable defined by p_E (>1/2) if $f(E)=1$ and $1-p_E$ (<1/2) if $f(E)=0$ (<1/2). Among these two possible limits the one that corresponds to a majority for the correct alternative happens with probability $P(f(E)=x|X=x)$. Hence, with probability 1, there is convergence

5. SUMMARY

On the basis of the Bayesian tree shown in diagram 2 above, we have developed a model of jury decision making where all jurors have the same evidence. We have suggested that the new model is more realistic than the classical Condorcet jury model. First, it captures the fact that in real world jury decisions the available evidence is limited, and that it is simply not possible to find arbitrarily many jurors who each have an *independent* signal about the true state of the world. Second, our model is more consistent with the common legal requirement that jurors must not obtain or use any evidence other than the one presented in the courtroom. This means that, *even if* – hypothetically – the jurors could each obtain an independent signal about the true state of the world, they would be required by law not to make use of such information.

Our model makes three key assumptions:

- According to the common signal assumption, the jurors' votes depend on the true state of the world only *through* the available evidence.
- According to the competence assumption, for each possible body of evidence e, each juror has a probability p_e greater than ½ of matching the *ideal* interpretation of the evidence e. On the homogeneous version of the competence assumption p_e is the same for all possible bodies of evidence e, whereas on the heterogeneous version p_e may depend on e.
- According to the independence assumption, the votes of different jurors are independent from each other *conditional on the evidence*.

Then:

- The probability of a majority decision that matches the true state of the world (given that state of the world) is typically less than, and at most equal to, the corresponding probability in the classical Condorcet jury model (assuming homogeneous competence in our model).
- As the number of jurors tends to infinity, the probability of a majority decision matching the true state of the world (given that state of the world) converges to the probability that the evidence is not misleading, i.e. to the probability that the ideal interpretation of the evidence matches the true state of the world. Unless the evidence is never misleading, that probability is strictly less than one.
- Our model reduces to the classical Condorcet jury model *if and only if* we assume *both* that the evidence is never misleading *and* that the competence parameter p_e is the same for *all* possible bodies of evidence (homogeneous competence). If we think that these assumptions are inadequate in the case of real world jury decisions, it follows that the classical Condorcet jury model, as it stands, fails to apply to real world jury decisions.

---

to a stable majority as the number of jurors increases, where this majority supports the correct alternative with the probability that the evidence "tells the truth".

APPENDIX

Proof of theorem 3.

First, we use the law of iterated expectations to write

$P(V=v|X=x) = Exp(P(V=v|E \text{ and } X=x)|X=x)$.

By (S) we have $P(V=v|E \text{ and } X=x) = P(V=v|E)$, so that we deduce

(*) $P(V=v|X=x) = Exp(P(V=v|E)|X=x)$.

By (C) and (I), conditional on E the votes $V_1, V_2, ..., V_n$ are independent and Bernoulli distributed with parameter p_E if $f(E)=1$ and $1-p_E$ if $f(E)=0$. Hence the sum V has a binomial distribution with first parameter n and second parameter p_E if $f(E)=1$ and $1-p_E$ if $f(E)=0$:

$$P(V=v|E) = \begin{cases} \binom{n}{v} p_E^{v}(1-p_E)^{n-v} & \text{if } f(E)=1 \\ \binom{n}{v} p_E^{n-v}(1-p_E)^{v} & \text{if } f(E)=0. \end{cases}$$

In other words,

$$P(V=v|E) = \binom{n}{v} p_E^{v}(1-p_E)^{n-v} 1_{\{f(E)=1\}} + \binom{n}{v} p_E^{n-v}(1-p_E)^{v\,v} 1_{\{f(E)=0\}},$$

where $1_{\{f(E)=1\}}$ and $1_{\{f(E)=0\}}$ are characteristic functions (1_A is the random variable defined as 1 if the event A holds and as 0 if it doesn't).

So, by (*) and the linearity of the (conditional) expectation operator $Exp(.|X=x)$,

$$P(V=v|X=x) = P(f(E)=1|X=x)\binom{n}{v}Exp(p_E^{v}(1-p_E)^{n-v}|f(E)=1 \text{ and } X=x)$$
$$+ P(f(E)=0|X=x)\binom{n}{v}Exp(p_E^{n-v}(1-p_E)^{v}|f(E)=0 \text{ and } X=x). \square$$

Proof of theorem 4.

(i) Let W be the number of jurors $i \in \{1, 2, ..., n\}$ such that $V_i = f(E)$.

We conditionalize on E. By (C) and (I), W is the sum of n independent Bernoulli variables with parrameter p_E. The weak law of large numbers implies that the average W/n converges in probability to p_E. Since $p_E > 1/2$, it follows that

$lim_{n\to\infty} P(W > n/2|E) = 1$.

Applying the (conditional) expectation operator on both sides (which corresponds to averaging with respect to E), we obtain

$Exp(lim_{n\to\infty}P(W>n/2|E) \mid X=x))=Exp(1 \mid X=x)=1$.

By the dominated convergence theorem, we can interchange the expectation operator with the limit operator on the left hand side, so that

$lim_{n\to\infty}\, Exp(P(W>n/2|E) \mid X=x) = 1$.

By (S) we can replace $P(W>n/2|E)$ by $P(W>n/2|E \text{ and } X=x)$. This leads to

$lim_{n\to\infty}\, Exp(P(W>n/2|E \text{ and } X=x) \mid X=x) = 1$,

and hence by the law of iterated expectations

$lim_{n\to\infty}\, P(W>n/2 \mid X=x) = 1$.

(ii) Using the weak law of large numbers in a similar way as in (i), it is possible to prove that the probability $P(V>n/2|E)=P(V/n>1/2|E)$ converges to 1 if f(E)=1 and to 0 if $f(E)=0$ (as n tends to infinity). Hence

(*) $lim_{n\to\infty}\, P(V>n/2 \mid E)=1_{\{f(E)=1\}}$,

where $1_{\{f(E)=1\}}$ is the random variable defined as 1 if $f(E)=1$ and as 0 if $f(E)=0$.

By the law of iterated expectations,

$P(V>n/2|X=1)=Exp(P(V>n/2|E$ and $X=1)|X=1)$, which by (S) simplifies to:

(**) $P(V>n/2|X=1)= Exp(P(V>n/2|E)|X=1)$.

Further, we have

$P(f(E)=1|X=1)=Exp(1_{\{f(E)=1\}}\, |X=1)=Exp(lim_{n\to\infty}\, P(V>n/2 \mid E)|X=1)$,

where the last step uses (*). We now interchange the expectation operator with the limit (by the dominated convergence theorem) and then use (**) to obtain

$P(f(E)=1|X=1)=lim_{n\to\infty}\, Exp(P(V>n/2 \mid E)|X=1)= lim_{n\to\infty} P(V>n/2|X=1)$.

As for the case $X=0$, it can be shown similarly that

$P(f(E)=0|X=0)=lim_{n\to\infty}\, P(V<n/2|X=0)$. □

ACKNOWLEDGEMENTS

A previous version of this paper was presented at the International Summer School on Philosophy and Probability, University of Konstanz, September 2002. We are grateful to Luc Bovens, Branden Fitelson, Daniel Rost and Jon Williamson for comments and discussion.

REFERENCES

Austen-Smith, D., and J. Banks (1996), Information Aggregation, Rationality, and the Condorcet Jury Theorem, *American Political Science Review* 90: 34-45.

Berend, D., and J. Paroush (1998), When is Condorcet's jury theorem valid, *Social Choice and Welfare* 15: 481-488.

Boland, P. J. (1989), Majority Systems and the Condorcet Jury Theorem, *Statistician* 38: 181-189.

Boland, P. J., F. Proschan and Y. L. Tong (1989), Modelling dependence in simple and indirect majority systems, *Journal of Applied Probability* 26: 81-88.

Estlund, D. (1994), Opinion leaders, independence and Condorcet's jury theorem, *Theory and Decision* 36: 131-162.

Grofman, B., G. Owen and S. L. Feld (1983), Thirteen theorems in search of the truth, *Theory and Decision* 15: 261-278.

Lahda, K. K. (1992), The Condorcet Jury Theorem, Free Speech, and Correlated Votes, *American Journal of Political Science* 36: 617-634.

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

Nitzan, S., and J. Paroush (1984), The significance of independent decisions in uncertain dichotomous choice situations, *Theory and Decision* 17: 47-60.

Young, H. P. (1988), Condorcet's theory of voting, *American Political Sciene Review* 82: 1231-1244.

Addresses for correspondence: F. Dietrich, Group on Philosophy, Probability and Modelling, Center for Junior Research Fellows, University of Konstanz, 78457 Konstanz, Germany; franz.dietrich@brasenose.oxford.ac.uk; C. List, Nuffield College, Oxford OX1 1NF, U.K.; christian.list@nuffield.oxford.ac.uk.