Condorcet Jury Theorem: The dependent case

Bezalel Peleg and Shmuel Zamir[1]

Center for the Study of Rationality
The Hebrew University of Jerusalem.

Abstract

We provide an extension of the Condorcet Theorem. Our model includes both the Nitzan-Paroush framework of "unequal competencies" and Ladha's model of "correlated voting by the jurors." We assume that the jurors behave "informatively"; that is, they do not make a strategic use of their information in voting. Formally, we consider a sequence of binary random variables $X = (X_1, X_2, ..., X_n, ...)$ with range in $\{0,1\}$ and a joint probability distribution P. The pair (X,P) is said to satisfy the *Condorcet Jury Theorem* (CJT) if $\lim_{n\to\infty} P\left(\Sigma_{i=1}^n X_i > \frac{n}{2}\right) = 1$. For a general (dependent) distribution P we provide necessary as well as sufficient conditions for the CJT that establishes the validity of the CJT for a domain that strictly (and naturally) includes the domain of independent jurors. In particular we provide a full characterization of the exchangeable distributions that satisfy the CJT. We exhibit a large family of distributions P with $\liminf_{n\to\infty} \frac{1}{n(n-1)} \Sigma_{i=1}^n \Sigma_{j\neq i} Cov(X_i, X_j) > 0$ that satisfy the CJT. We do that by "interlacing" carefully selected pairs (X,P) and (X',P'). We then proceed to project the distributions P on the planes $(\underline{p}, \underline{y})$, and $(\underline{p}, \underline{y}^*)$ where $\underline{p} = \liminf_{n\to\infty} \overline{p}_n$, $\underline{y} = \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2$, $\underline{y}^* = \liminf_{n\to\infty} E|\overline{X}_n - \overline{p}_n|$, $\overline{p}_n = (p_1 + p_2, ... + p_n)/n$, and $\overline{X}_n = (X_1 + X_2, ... + X_n)/n$. We determine all feasible points in each of these planes. Quite surprisingly, many important results on the possibility of the CJT are obtained by analyzing various regions of the feasible set in these planes.

In the space $\mathscr{P}$ of all probability distributions on $S_p = \{0,1\}^\infty$, let $\mathscr{P}_1$ be the set of all probability distributions in $\mathscr{P}$ that satisfy the CJT and let $\mathscr{P}_2 = \mathscr{P} \setminus \mathscr{P}_1$ be the set of all probability distributions in $\mathscr{P}$ that do not satisfy the CJT. We prove that both $\mathscr{P}_1$ and $\mathscr{P}_2$ are convex sets and that $\mathscr{P}_2$ is dense in $\mathscr{P}$ (in the weak topology). Using an appropriate separation theorem we then provide an affine functional that separates these two sets.

[1]We thank Marco Scarsini and Yosi Rinott for drawing our attention to de Finetti's theorem.

Introduction

The simplest way to present our problem is by quoting Condorcet's classic result (see Young (1997)):

Theorem 1. *(CJT–Condorcet 1785) Let n voters (n odd) choose between two alternatives that have equal likelihood of being correct a priori. Assume that voters make their judgements independently and that each has the same probability p of being correct ($\frac{1}{2} < p < 1$). Then, the probability that the group makes the correct judgement using simple majority rule is*

$$\sum_{h=(n+1)/2}^{n} [n!/h!(n-h)!] p^h (1-p)^{n-h}$$

which approaches 1 *as n becomes large.*

We generalize Condorcet's model by presenting it as a *game with incomplete information* in the following way: Let $I = \{1, 2, \ldots, n\}$ be a set of jurors and let D be the defendant. There are two *states of nature*: g – in which D is guilty, and z – in which D is innocent. Thus the set of states of nature is $S = \{g, z\}$. Each juror has an action set A with two actions: $A = \{c, a\}$. The action c is to *convict* D. The action a is to *acquit* D. Before the voting, each juror i gets a private random signal $t^i \in T^i := \{t^i_g, t^i_z\}$. In the terminology of games with incomplete information, T^i is the *type set* of juror i. The interpretation is that juror i of type t^i_g thinks that D is guilty while juror i of type t^i_z thinks that D is innocent. The signals of the jurors may be dependent and may also depend on the the state of nature. In our model the jurors act "*informatively*" (not "strategically"); that is, the strategy of juror i is $\sigma^i : T^i \to A$ given by $\sigma^i(t^i_g) = c$ and $\sigma^i(t^i_z) = a$. The definition of informative voting is due to Austen-Smith and Banks (1996), who question the validity of the CJT in a strategic framework. Informative voting was, and is still, assumed in the vast majority of the literature on the CJT, mainly because it is implied by the original Condorcet assumptions. More precisely, assume, as Condorcet did, that $P(g) = P(z) = 1/2$ and that each juror is more likely to receive the "correct" signal (that is, $P(t^i_g|g) = P(t^i_z|z) = p > 1/2$); then the strategy of voting informatively maximizes the probability of voting correctly, among all four pure voting strategies. Following Austen-Smith and Banks, strategic voting and Nash Equilibrium were studied by Wit (1998), Myerson (1998), and recently by Laslier and Weibull (2008), who discuss the assumption on preferences and beliefs under which sincere voting is a Nash equilibrium in a general deterministic majoritarian voting rule. As we said before, in this work we do assume informative voting and leave strategic considerations and equilibrium concepts for the next phase of our research. The action taken by a finite society of jurors $\{1, \ldots, n\}$ (i.e. the jury verdict) is determined by a simple majority (with some tie-breaking rule, e.g., by coin tossing). We are interested in the probability that the (finite) jury will reach the correct decision. Again in the style of games with incomplete information let $\Omega_n = S \times T^1 \times, \dots, \times T^n$ be the set of *states of the world*. A state of the world consists of the state of nature and a list of the types of all jurors. Denote by $p^{(n)}$ the probability distribution on Ω_n. This is a joint probability distribution on the state of nature and the signals of the jurors. For each juror i let the random variable $X_i : S \times T^i \to \{0,1\}$ be the indicator of his correct voting, i.e., $X_i(g, t_g^i) = X_i(z, t_z^i) = 1$ and $X_i(g, t_z^i) = X_i(z, t_g^i) = 0$. The probability distribution $p^{(n)}$ on Ω_n induces a joint probability distribution on the the vector $X = (X_1, \dots, X_n)$, which we denote also by $p^{(n)}$. If n is odd, then the probability that the jury reaches a correct decision is

$$p^{(n)} \left(\sum_{i=1}^{n} X_i > \frac{n}{2} \right)$$

.

Figure 1 illustrates our construction in the case $n = 2$. In this example, according to $p^{(2)}$ the state of nature is chosen with unequal probabilities for the two states: $P(g) = 1/4$ and $P(z) = 3/4$ and then the types of the two jurors are chosen according to a joint probability distribution that depends on the state of nature.

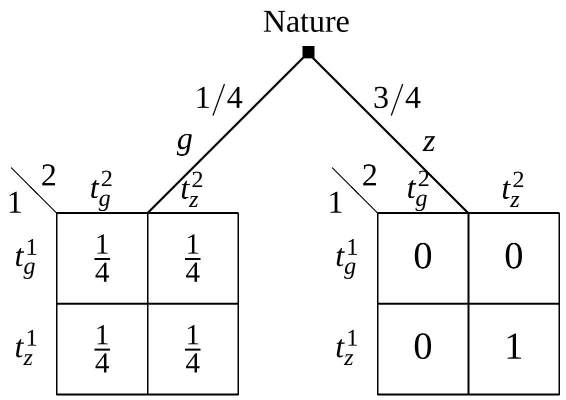

Figure 1 The probability distribution $p^{(2)}$.

Guided by Condorcet, we are looking for limit theorems as the the size of the jury increases. Formally, as n goes to infinity we obtain an increasing sequence of "worlds", $(\Omega_n)_{n=1}^{\infty}$, such that for all n, the projection of Ω_{n+1} on Ω_n is the whole Ω_n. The corresponding sequence of probability distributions is $(p^{(n)})_{n=1}^{\infty}$ and we assume that for every n, the marginal distribution of $p^{(n+1)}$ on Ω_n is $p^{(n)}$. It follows from the Kolmogorov extension theorem (see Loeve (1963), p. 93) that this defines a unique probability measure P on the (projective, or *inverse*) limit

$$\Omega = \lim_{\infty \leftarrow n} \Omega_n = S \times T^1 \times \dots \times T^n \dots$$

such that, for all n, the marginal distribution of P on Ω_n is $p^{(n)}$.

In this paper we address the the following problem: Which probability measures P derived in this manner satisfy the *Condorcet Jury Theorem* (CJT); that is, Which probability measures P satisfy

$$\lim_{n\to\infty} P\left(\Sigma_{i=1}^{n} X_i > \frac{n}{2}\right) = 1.$$

As far as we know, the only existing result on this general problem is that of Berend and Paroush (1998), which deals only with independent jurors.

Rather than working with the space Ω and its probability measure P, it will be more convenient to work with the infinite sequence of binary random variables $X = (X_1, X_2, ..., X_n, ...)$ (the indicators of "correct voting") and the induced probability measure on it, which we shall denote also by P. Since the pair (X,P) is uniquely determined by (Ω,P) , in considering all pairs (X,P) we cover all pairs (Ω,P).

We provide a full characterization of the exchangeable sequences $X = (X_1, X_2, ..., X_n, ...)$ that satisfy the CJT. For a general (dependent) distribution P, not necessarily exchangeable, we provide necessary as well as sufficient conditions for the CJT. We exhibit a large family of distributions P with $\liminf_{n\to\infty} \frac{1}{n(n-1)} \Sigma_{i=1}^{n} \Sigma_{j\neq i} Cov(X_i,X_j) > 0$ that satisfy the CJT.

In the space $\mathscr{P}$ of all probability distributions on $S_p = \{0,1\}^\infty$, let $\mathscr{P}_1$ be the set of all probability distributions in $\mathscr{P}$ that satisfy the CJT and let $\mathscr{P}_2 = \mathscr{P} \setminus \mathscr{P}_1$ be the set of all probability distributions in $\mathscr{P}$ that do not satisfy the CJT. We prove that both $\mathscr{P}_1$ and $\mathscr{P}_2$ are convex sets and that $\mathscr{P}_2$ is dense in $\mathscr{P}$ (in the weak topology). Using an appropriate separation theorem we then provide an affine functional that separates these two sets.

1 Sufficient conditions

Let $X = (X_1, X_2, ..., X_n, ...)$ be a sequence of binary random variables with range in $\{0,1\}$ and with joint probability distribution P. The sequence X is said to satisfy the *Condorcet Jury Theorem* (CJT) if

$$\lim_{n\to\infty} P\left(\Sigma_{i=1}^{n} X_i > \frac{n}{2}\right) = 1 \tag{1}$$

We shall investigate necessary as well as sufficient conditions for CJT.

Given a sequence of random binary variables $X = (X_1, X_2, ..., X_n, ...)$ with joint distribution P denote $p_i = E(X_i)$, $Var(X_i) = E(X_i - p_i)^2$ and $Cov(X_i, X_j) = E[(X_i - p_i)(X_j - p_j)]$, for $i \neq j$, where E denotes, as usual, the expectation operator. Also let $\overline{p}_n = (p_1 + p_2, ... + p_n)/n$ and $\overline{X}_n = (X_1 + X_2, ... + X_n)/n$.

Our first result provides a sufficient condition for CJT:

Theorem 2. *Assume that $\Sigma_{i=1}^{n} p_i > \frac{n}{2}$ for all $n > N_0$ and*

$$\lim_{n\to\infty} \frac{E(\overline{X}_n - \overline{p}_n)^2}{(\overline{p}_n - \frac{1}{2})^2} = 0, \tag{2}$$

or equivalently assume that

$$\lim_{n\to\infty} \frac{\overline{p}_n - \frac{1}{2}}{\sqrt{E(\overline{X}_n - \overline{p}_n)^2}} = \infty; \tag{3}$$

then the CJT is satisfied.

Proof.

$$\begin{aligned} P\left(\Sigma_{i=1}^n X_i \le \frac{n}{2}\right) &= P\left(-\Sigma_{i=1}^n X_i \ge -\frac{n}{2}\right) \\ &= P\left(\Sigma_{i=1}^n p_i - \Sigma_{i=1}^n X_i \ge \Sigma_{i=1}^n p_i - \frac{n}{2}\right) \\ &\le P\left(|\Sigma_{i=1}^n p_i - \Sigma_{i=1}^n X_i| \ge \Sigma_{i=1}^n p_i - \frac{n}{2}\right) \end{aligned}$$

By Chebyshev's inequality (assuming $\Sigma_{i=1}^n p_i > \frac{n}{2}$) we have

$$P\left(|\Sigma_{i=1}^n p_i - \Sigma_{i=1}^n X_i| \ge \Sigma_{i=1}^n p_i - \frac{n}{2}\right) \le \frac{E\left(\Sigma_{i=1}^n X_i - \Sigma_{i=1}^n p_i\right)^2}{\left(\Sigma_{i=1}^n p_i - \frac{n}{2}\right)^2} = \frac{E(\overline{X}_n - \overline{p}_n)^2}{(\overline{p}_n - \frac{1}{2})^2}$$

As this last term tends to zero by (2), the CJT (1) then follows. □

Corollary 3. *If $\Sigma_{i=1}^n \Sigma_{j\neq i} Cov(X_i, X_j) \le 0$ for $n > N_0$ (in particular if $Cov(X_i, X_j) \le 0$ for all $i \neq j$) and $\lim_{n\to\infty} \sqrt{n}(\overline{p}_n - \frac{1}{2}) = \infty$, then the CJT is satisfied.*

Proof. Since the variance of a binary random variable X with mean p is $p(1-p) \le 1/4$ we have for $n > N_0$,

$$\begin{aligned} 0 \le E(\overline{X}_n - \overline{p}_n)^2 &= \frac{1}{n^2} E\left(\Sigma_{i=1}^n (X_i - p_i)\right)^2 \\ &= \frac{1}{n^2}\left(\Sigma_{i=1}^n Var(X_i) + \Sigma_{i=1}^n \Sigma_{j\neq i} Cov(X_i, X_j)\right) \le \frac{1}{4n} \end{aligned}$$

Therefore if $\lim_{n\to\infty} \sqrt{n}(\overline{p}_n - \frac{1}{2}) = \infty$, then

$$0 \le \lim_{n\to\infty} \frac{E(\overline{X}_n - \overline{p}_n)^2}{(\overline{p}_n - \frac{1}{2})^2} \le \lim_{n\to\infty} \frac{1}{4n(\overline{p}_n - \frac{1}{2})^2} = 0$$

□

Remark 4. *Note that under the condition of corollary 3, namely, for bounded random variable with all covariances being non-positive, the (weak) law of large numbers (LLN) holds for arbitrarily dependent variables (see, e.g., Feller (1957) volume I, exercise 9, p. 262). This is not implied by corollary 3 since, as we shall see later, the CJT, strictly speaking, is not a law of large numbers. In particular, CJT does not imply LLN and LLN does not imply CJT.*

Remark 5. *When $X_1, X_2, \ldots, X_n, \ldots$ are independent, then under mild conditions $\lim_{n\to\infty}\sqrt{n}(\overline{p}_n - \frac{1}{2}) = \infty$ is a necessary and sufficient condition for CJT (see Berend and Paroush (1998)).*

Remark 6. *The sequence $X = (X_1, X_2, \ldots, X_n, \ldots)$ of i.i.d. variables with $P(X_i = 1) = P(X_i = 0) = 1/2$ satisfies the LLN but does not satisfy the CJT, since it does not satisfy Berend and Paroush's necessary and sufficient condition; therefore LLN does not imply CJT.*

Given a sequence $X = (X_1, X_2, \ldots, X_n, \ldots)$ of binary random variables with a joint probability distribution P, we define the following parameters of (X, P):

$$\underline{p} := \liminf_{n\to\infty} \overline{p}_n \tag{4}$$
$$\overline{p} := \limsup_{n\to\infty} \overline{p}_n \tag{5}$$
$$\underline{y} := \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 \tag{6}$$
$$\overline{y} := \limsup_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 \tag{7}$$
$$\underline{y}^* := \liminf_{n\to\infty} E|\overline{X}_n - \overline{p}_n| \tag{8}$$
$$\overline{y}^* := \limsup_{n\to\infty} E|\overline{X}_n - \overline{p}_n| \tag{9}$$

We first observe the following:

Remark 7. *If $\underline{p} > 1/2$ and $\overline{y} = 0$ then the CJT is satisfied.*

Proof. As $E(\overline{X}_n - \overline{p}_n)^2 \geq 0$, if $\overline{y} = 0$ then $\lim_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 = 0$. Since $\underline{p} > 1/2$, there exists n_0 such that $\overline{p}_n > (1/2 + \underline{p})/2$ for all $n > n_0$. The result then follows by Theorem 2. □

2 Necessary conditions using the L_1-norm

Given a sequence $X = (X_1, X_2, \ldots, X_n, \ldots)$ of binary random variables with a joint probability distribution P, if $\overline{y} > 0$, then we cannot use Theorem 2 to conclude CJT.

To derive necessary conditions for the CJT, we first have:

Proposition 8. *If the CJT holds then $\underline{p} \geq \frac{1}{2}$.*

Proof. Define a sequence of events $(B_n)_{n=1}^{\infty}$ by $B_n = \{\omega | \overline{X}_n(\omega) - 1/2 \geq 0\}$. Since the CJT holds, $\lim_{n\to\infty} P\left(\Sigma_{i=1}^n X_i > \frac{n}{2}\right) = 1$ and hence $\lim_{n\to\infty} P(B_n) = 1$. Since

$$\overline{p}_n - \frac{1}{2} = E(\overline{X}_n - \frac{1}{2})) \geq -\frac{1}{2} P(\Omega \backslash B_n),$$

taking the lim inf, the right-hand side tends to zero and we obtain that $\liminf_{n\to\infty} p_n = \underline{p} \geq \frac{1}{2}$. □

We shall first consider a stronger violation of Theorem 2 than $\bar{y} > 0$; namely, assume that $\underline{y} > 0$. We shall prove that in this case, there is a range of distributions P for which the CJT is false.

First we notice that for $-1 \leq x \leq 1$, $|x| \geq x^2$. Hence $E|\overline{X}_n - \overline{p}_n| \geq E(\overline{X}_n - \overline{p}_n)^2$ for all n and thus $\underline{y} > 0$ implies that $\underline{y}^* > 0$

We are now ready to state our first impossibility theorem, which can be readily translated into a necessary condition.

Theorem 9. *Given a sequence $X = (X_1, X_2, ..., X_n, ...)$ of binary random variables with joint probability distribution P, if $\underline{p} < \frac{1}{2} + \frac{\underline{y}^*}{2}$, then the (X,P) violates the CJT.*

Proof. If $\underline{y}^* = 0$, then the CJT is violated by Proposition 8. Assume then that $\underline{y}^* > 0$ and choose $\tilde{y}$ such that $0 < \tilde{y} < \underline{y}^*$ and $2t := \frac{\tilde{y}}{2} + \frac{1}{2} - \underline{p} > 0$. First we notice that, since $E(\overline{X}_n - \overline{p}_n) = 0$, we have $E\max(0, \overline{p}_n - \overline{X}_n) = E\max(0, \overline{X}_n - \overline{p}_n)$, and thus, since $\underline{y}^* > 0$, we have

$$E\max(0, \overline{p}_n - \overline{X}_n) > \frac{\tilde{y}}{2} \text{ for } n > \overline{n}. \tag{10}$$

If (Ω, P) is the probability space on which the sequence X is defined, for $n > \overline{n}$ define the events

$$B_n = \{\omega | \overline{p}_n - \overline{X}_n(\omega) \geq \max(0, \frac{\tilde{y}}{2} - t)\} \tag{11}$$

By (10) and (11), $P(B_n) > q > 0$ for some q and

$$\overline{p}_n - \overline{X}_n(\omega) \geq \frac{\tilde{y}}{2} - t \text{ for } \omega \in B_n,\ n > \overline{n}. \tag{12}$$

Choose now a subsequence $(n_k)_{k=1}^{\infty}$ such that

$$\overline{p}_{n_k} < \frac{\tilde{y}}{2} + \frac{1}{2} - t = \underline{p} + t,\ \ k = 1, 2, ... \tag{13}$$

By (12) and (13), for all $\omega \in B_{n_k}$ we have,

$$\overline{X}_{n_k}(\omega) \leq \overline{p}_{n_k} - \frac{\tilde{y}}{2} + t < \frac{1}{2},$$

and thus $P(\overline{X}_{n_k} > \frac{1}{2}) \leq 1 - q < 1$, which implies that P violates the CJT. □

Corollary 10. *If $\liminf_{n\to\infty} \overline{p}_n \leq \frac{1}{2}$ and $\liminf_{n\to\infty} E|\overline{X}_n - \overline{p}_n| > 0$, then P violates the CJT.*

3 Necessary conditions using the L_2-norm

Let $X = (X_1, X_2, ..., X_n, ...)$ be a sequence of binary random variables with a joint probability distribution P. In this section we take a closer look at the relationship between the parameters $\underline{y}$ and $\underline{y}^*$(see (7) and (9)). We first notice that $\underline{y} > 0$ if and only if $\underline{y}^* > 0$. Next we notice that $\overline{p}_n \geq \frac{1}{2}$ for $n > \overline{n}$ implies that $\overline{X}_n - \overline{p}_n \leq \frac{1}{2}$ for $n > \overline{n}$. Thus, by corollary 10, if $\underline{y} > 0$ and the CJT is satisfied then $\max(0, \overline{X}_n - \overline{p}_n) \leq \frac{1}{2}$ for $n > \overline{n}$. Finally we observe the following lemma, whose proof is straightforward:

Lemma 11. *If* $\limsup_{n\to\infty} P\{\omega | \overline{p}_n - \overline{X}_n(\omega) \geq \overline{p}_n/2\} > 0$, *then the* CJT *is violated.*

We now use the previous discussion to prove the following theorem:

Theorem 12. *If (i)* $\liminf_{n\to\infty} \overline{p}_n > \frac{1}{2}$ *and (ii)* $\liminf_{n\to\infty} P(\overline{X}_n > \overline{p}_n/2) = 1$, *then* $\underline{y}^* \geq 2\underline{y}$.

Proof. Condition (i) implies that $\overline{p}_n > \frac{1}{2}$ for all $n > N_0$ for some N_0. This implies that $\lim_{n\to\infty} P(\overline{X}_n - \overline{p}_n < \frac{1}{2}) = 1$. Also, (ii) implies that $\lim_{n\to\infty} P(\overline{p}_n - \overline{X}_n \leq \frac{1}{2}) = 1$, and thus

$$\lim_{n\to\infty} P(-\frac{1}{2} \leq \overline{X}_n - \overline{p}_n \leq \frac{1}{2}) = 1 \tag{14}$$

Define the events $B_n = \{\omega | -\frac{1}{2} \leq \overline{X}_n(\omega) - \overline{p}_n \leq \frac{1}{2}\}$; then by (14)

$$\liminf_{n\to\infty} \int_{B_n} (\overline{X}_n - \overline{p}_n)^2 dP = \underline{y} \tag{15}$$

and

$$\liminf_{n\to\infty} \int_{B_n} |\overline{X}_n - \overline{p}_n| dP = \underline{y}^*. \tag{16}$$

Since any $u \in [-\frac{1}{2}, \frac{1}{2}]$ satisfies $|u| \geq 2u^2$, it follows from (15) and (16) that $\underline{y}^* \geq 2\underline{y}$. □

Corollary 13. *Let* $\underline{p} = \liminf_{n\to\infty} \overline{p}_n$ *and* $\underline{y} = \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2$. *Then if* $\underline{p} < \frac{1}{2} + \underline{y}$ *then* P *does not satisfy the* CJT.

Proof. Assume that $\underline{p} < \frac{1}{2} + \underline{y}$. If $\underline{y} = 0$, then CJT is not satisfied by Proposition 8. Hence we may thus assume that $\underline{y} > 0$, which also implies that $\underline{y}^* > 0$. Thus, if $\liminf_{n\to\infty} \overline{p}_n \leq \frac{1}{2}$ then CJT fails by Corollary 10. Assume then that $\liminf_{n\to\infty} \overline{p}_n > \frac{1}{2}$. By Lemma 11 we may also assume that $\liminf_{n\to\infty} P\left(\overline{X}_n > \overline{p}_n/2\right) = 1$ and thus by Theorem 12 we have $\underline{y}^* \geq 2\underline{y}$ and hence $\underline{p} < \frac{1}{2} + \underline{y} \leq \frac{1}{2} + \frac{\underline{y}^*}{2}$ and the CJT fails by Theorem 9. □

4 Dual Conditions

A careful reading of Sections 2 and 3 reveals that it is possible to obtain "dual" results to Theorems 9 and 12 and Corollary 13 by replacing "lim inf" by "lim sup". More precisely, for a sequence $X = (X_1, X_2, \dots, X_n, \dots)$ of binary random variables with joint probability distribution P, we let $\overline{p} = \limsup_{n\to\infty} \overline{p}_n$ and $\overline{y}^* = \limsup_{n\to\infty} E|\overline{X}_n - \overline{p}_n|$, and we have:

Theorem 14. *If* $\overline{p} < \frac{1}{2} + \frac{\overline{y}^*}{2}$*, then the* (X,P) *violates the* CJT*.*

Proof. As we saw in the proof of Corollary 13, we may assume that $\liminf_{n\to\infty} \overline{p}_n \geq \frac{1}{2}$ and hence also

$$\overline{p} = \limsup_{n\to\infty} \overline{p}_n \geq \liminf_{n\to\infty} \overline{p}_n \geq \frac{1}{2},$$

and hence $\overline{y}^* > 0$. Choose $\tilde{y}$ such that $0 < \tilde{y} < \overline{y}^*$ and $2t = \frac{\tilde{y}}{2} + \frac{1}{2} - \overline{p} > 0$. Let $(X_{n_k})_{k=1}^{\infty}$ be a subsequence of X such that

$$\lim_{k\to\infty} E|\overline{X}_{n_k} - \overline{p}_{n_k}| = \overline{y}^*.$$

As in (10) we get

$$E \max(0, \overline{p}_{n_k} - \overline{X}_{n_k}) > \frac{\tilde{y}}{2} \;\text{ for }\; k > \overline{k}. \tag{17}$$

Define the events $(B_{n_k})_{k=1}^{\infty}$ by

$$B_{n_k} = \{\omega | \overline{p}_{n_k} - \overline{X}_{n_k}(\omega) \geq \frac{\tilde{y}}{2} - t\}. \tag{18}$$

By (17) and (18), $P(B_{n_k}) > q$ for some $q > 0$ and

$$\overline{p}_{n_k} - \overline{X}_{n_k}(\omega) \geq \frac{\tilde{y}}{2} - t \;\text{ for }\; \omega \in B_{n_k} \;\text{ and }\; k > \overline{\overline{k}}. \tag{19}$$

Now

$$\limsup_{n\to\infty} \overline{p}_n = \overline{p} < \frac{\tilde{y}}{2} + \frac{1}{2} - t. \tag{20}$$

Thus, for n sufficiently large $\overline{p}_n < \frac{\tilde{y}}{2} + \frac{1}{2} - t$. Hence, for k sufficiently large and all $\omega \in B_{n_k}$,

$$\overline{X}_{n_k}(\omega) \leq \overline{p}_{n_k} - \frac{\tilde{y}}{2} + t < \frac{1}{2}. \tag{21}$$

Therefore $P(\overline{X}_{n_k} > \frac{1}{2}) \leq 1 - q < 1$ for sufficiently large k in violation of the CJT. □

Similarly we have the "dual" results to those of Theorem 12 and Corollary 13:

Theorem 15. *If (i)* $\liminf_{n\to\infty} \overline{p}_n > \frac{1}{2}$ *and*
(ii) $\liminf_{n\to\infty} P(\overline{X}_n > \overline{p}_n/2) = 1$*, then* $\overline{y}^* \geq 2\overline{y}$*.*

Corollary 16. *If* $\overline{p} < \frac{1}{2} + \overline{y}$ *then* P *does not satisfy the* CJT*.*

The proofs, which are similar respectively to the proofs of Theorem 12 and Corollary 13, are omitted.

5 Existence of distributions satisfying the CJT

In this section we address the issue of the existence of distributions that satisfy the CJT. For that, let us first clarify the relationship between the CJT and law of large numbers which, at first sight, look rather similar. Recall that an infinite sequence of binary random variables $X = (X_1, X_2, \ldots, X_n, \ldots)$ with a joint probability distribution P satisfies the (weak) law of large numbers (LLN) if (in our notations):

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P\left(|\overline{X}_n - \overline{p}_n| < \varepsilon\right) = 1 \tag{22}$$

while it satisfies the Condorcet Jury Theorem (CJT) if:

$$\lim_{n\to\infty} P\left(\overline{X}_n > \frac{1}{2}\right) = 1 \tag{23}$$

Since by Proposition 8, the condition $\underline{p} \geq \frac{1}{2}$ is necessary for the validity of the CJT, let us check the relationship between the LLN and the CJT in this region. Our first observation is:

Proposition 17. *For a sequence $X = (X_1, X_2, \ldots, X_n, \ldots)$ with probability distribution P satisfying $\underline{p} > \frac{1}{2}$, if the LLN holds then the CJT also holds.*

Proof. Let $\underline{p} = 1/2 + 3\delta$ for some $\delta > 0$ and let N_0 be such that $\overline{p}_n > 1/2 + 2\delta$ for all $n > N_0$; then for all $n > N_0$ we have

$$P\left(\overline{X}_n > \frac{1}{2}\right) \geq P\left(\overline{X}_n \geq \frac{1}{2} + \delta\right) \geq P\left(|\overline{X}_n - \overline{p}_n| < \delta\right)$$

Since the last expression tends to 1 as $n \to \infty$, the first expression does too, and hence the CJT holds. □

Remark 18. *The statement of Proposition 17 does not hold for $\underline{p} = \frac{1}{2}$. Indeed, the sequence $X = (X_1, X_2, \ldots, X_n, \ldots)$ of i.i.d. variables with $P(X_i = 1) = P(X_i = 0) = 1/2$ satisfies the LLN but does not satisfy the CJT (see Remark 6).*

Unfortunately, Proposition 17 is of little use to us. This is due to the following fact:

Proposition 19. *If the random variables of the sequence $X = (X_1, X_2, \ldots, X_n, \ldots)$ are uniformly bounded then the condition*

$$\lim_{n\to\infty} E\left(\overline{X}_n - \overline{p}_n\right)^2 = 0$$

is a necessary condition for LLN to hold.

The proof is elementary and can be found, e.g., in Uspensky (1937), page 185. For the sake of completeness it is provided here in the Appendix.

It follows thus from Proposition 19 that LLN cannot hold when $\underline{y} > 0$ and thus we cannot use Proposition 17 to establish distributions in this region that satisfy the CJT. Nevertheless, we shall exhibit a rather large family of distributions P with $\underline{y} > 0$ (and $\underline{p} > 1/2$) for which the CJT holds. Our main result is the following:

Theorem 20. *Let $t \in [0, \frac{1}{2}]$. If F is a distribution with parameters $(\underline{p}, \underline{y}^*)$, then there exists a distribution H with parameters $\underline{\tilde{p}} = 1 - t + t\underline{p}$ and $\underline{\tilde{y}}^* = t\underline{y}^*$ that satisfy the CJT.*

Proof. To illustrate the idea of the proof we first prove (somewhat informally) the case $t = 1/2$. Let $X = (X_1, X_2, \ldots, X_n, \ldots)$ be a sequence of binary random variables with a joint probability distribution F. Let G be the distribution of the sequence $Y = (Y_1, Y_2, \ldots, Y_n, \ldots)$, where $EY_n = 1$ for all n (that is, $Y_1 = Y_2 = \ldots Y_n = \ldots$ and $P(Y_i = 1) = 1\ \forall i$). Consider now the following "interlacing" of the two sequences X and Y:

$$Z = (Y_1, Y_2, X_1, Y_3, X_2, Y_4, X_3, \ldots, Y_n, X_{n-1}, Y_{n+1}, X_n \ldots),$$

and let the probability distribution H of Z be the product distribution $H = F \times G$. It is verified by straightforward computation that the parameters of the distribution H are in accordance with the theorem for $t = \frac{1}{2}$, namely, $\underline{\tilde{p}} = \frac{1}{2} + \frac{1}{2}\underline{p}$ and $\underline{\tilde{y}}^* = \frac{1}{2}\underline{y}^*$. Finally, as each initial segment of voters in Z contains a majority of Y_i's (thus with all values 1), the distribution H satisfies the CJT, completing the proof for $t = \frac{1}{2}$.

The proof for a general $t \in [0, 1/2)$ follows the same lines: We construct the sequence Z so that any finite initial segment of n variables, includes "about, but not more than" the initial tn segment of the X sequence, and the rest is filled with the constant Y_i variables. This will imply that the CJT is satisfied.

Formally, for any real $x \geq 0$ let $\lfloor x \rfloor$ be the largest integer less than or equal to x and let $\lceil x \rceil$ be smallest integer greater than or equal to x. Note that for any n and any $0 \leq t \leq 1$ we have $\lfloor tn \rfloor + \lceil (1-t)n \rceil = n$; thus, one and only one of the following holds:

(i) $\lfloor tn \rfloor < \lfloor t(n+1) \rfloor$ or

(ii) $\lceil (1-t)n \rceil < \lceil (1-t)(n+1) \rceil$

From the given sequence X and the above-defined sequence Y (of constant 1 variables) we define now the sequence $Z = (Z_1, Z_2, \ldots, Z_n, \ldots)$ as follows: $Z_1 = Y_1$ and for any $n \geq 2$, let $Z_n = X_{\lfloor t(n+1) \rfloor}$ if (i) holds and $Z_n = Y_{\lceil (1-t)(n+1) \rceil}$ if (ii) holds. This inductive construction guarantees that for all n, the sequence contains $\lfloor tn \rfloor$ X_i coordinates and $\lceil (1-t)n \rceil$ Y_i coordinates. The probability distribution H is the product distribution $F \times G$. The fact that (Z, H) satisfies the CJT follows from:

$$\lceil (1-t)n \rceil \geq (1-t)n > tn \geq \lfloor tn \rfloor,$$

and finally $\underline{\tilde{p}} = 1 - t + t\underline{p}$ and $\underline{\tilde{y}}^* = t\underline{y}^*$ is verified by straightforward computation. □

Remark 21. *The "interlacing" of the two sequences X and Y described in the proof of Theorem 20 may be defined for any $t \in [0, 1]$. We were specifically interested in $t \in [0, 1/2]$ since this guarantees the CJT.*

6 Feasibility considerations

The conditions developed so far for a sequence $X = (X_1, X_2, ..., X_n, ...)$ with joint probability distribution P to satisfy the CJT involved only the parameters $\underline{p}, \overline{p}, \underline{y}, \overline{y}, \underline{y}^*$, and $\overline{y}^*$. In this section we pursue our characterization in the space of these parameters. We shall look at the distributions in two different spaces: the space of points $(\underline{p}, \underline{y}^*)$, which we call the L_1 space, and the space $(\underline{p}, \underline{y})$, which we call the L_2 space.

6.1 Feasibility and characterization in L_1

With the pair (X, P) we associate the point $(\underline{p}, \underline{y}^*)$ in the Euclidian plane $\mathbb{R}^2$. It follows immediately that $0 \leq \underline{p} \leq 1$. We claim that $\underline{y}^* \leq 2\underline{p}(1 - \underline{p})$ holds for all distributions P. To see that, we first observe that $E|X_i - p_i| = 2p_i(1 - p_i)$; hence

$$E|\overline{X}_n - \overline{p}_n| = \frac{1}{n} E|\sum_{i=1}^{n}(X_i - p_i)| \leq \frac{1}{n} E\left(\sum_{i=1}^{n}|X_i - p_i|\right) \leq \frac{2}{n}\sum_{i=1}^{n} p_i(1 - p_i).$$

The function $\sum_{i=1}^{n} p_i(1 - p_i)$ is (strictly) concave. Hence

$$E|\overline{X}_n - \overline{p}_n| \leq 2\sum_{i=1}^{n}\frac{1}{n} p_i(1 - p_i) \leq 2\overline{p}_n(1 - \overline{p}_n). \tag{24}$$

Finally, let $\underline{p} = \lim_{k\to\infty} \overline{p}_{n_k}$; then

$$\underline{y}^* = \liminf_{n\to\infty} E|\overline{X}_n - \overline{p}_n| \leq \liminf_{k\to\infty} E|\overline{X}_{n_k} - \overline{p}_{n_k}| \leq 2\lim_{k\to\infty} \overline{p}_{n_k}(1 - \overline{p}_{n_k}) = 2\underline{p}(1 - \underline{p}).$$

The second inequality is due to (24).

Thus, if (u, w) denote a point in $\mathbb{R}^2$, then any feasible pair $(\underline{p}, \underline{y}^*)$ is in the region

$$FE_1 = \{(u, w) | 0 \leq u \leq 1,\ \ 0 \leq w \leq 2u(1 - u)\}. \tag{25}$$

We shall now prove that all points in this region are feasible; that is, any point in FE_1 is attainable as a pair $(\underline{p}, \underline{y}^*)$ of some distribution P. Then we shall indicate the sub-region of FE_1 where the CJT may hold. We first observe that any point $(u_0, w_0) \in FE_1$ on the parabola $w = 2u(1 - u)$, for $0 \leq u \leq 1$, is feasible. In fact such (u_0, w_0) is attainable by the sequence $X = (X_1, X_2, ..., X_n, ...)$ with identical variables X_i, $X_1 = X_2 = ... = X_n ...$, and $EX_1 = u_0$ (clearly $\underline{p} = u_0$, and $\underline{y}^* = 2u_0(1 - u_0)$ follows from the dependence and from $E|X_i - p_i| = 2p_i(1 - p_i) = 2u_0(1 - u_0)$).

Let again (u_0, w_0) be a point on the parabola, which is thus attainable. Assume that they are the parameters $(\underline{p}, \underline{y}^*)$ of the pair (X, F). Let (Y, G) be the pair (of constant variables) described in the proof of Theorem 20 and let $t \in [0, 1]$. By Remark 21 the t-interlacing of (X, F) and (Y, G) can be constructed to yield a distribution with parameters $\underline{\tilde{p}} = t\underline{p} + (1 - t)$ and $\underline{\tilde{y}}^* = t\underline{y}^*$ (see the proof of Theorem 20). Thus, the line segment defined by $\tilde{u} = tu_0 + (1 - t)$ and $\tilde{w} = tw_0$ for $0 \leq t \leq 1$, connecting (u_0, w_0) to $(1, 0)$,

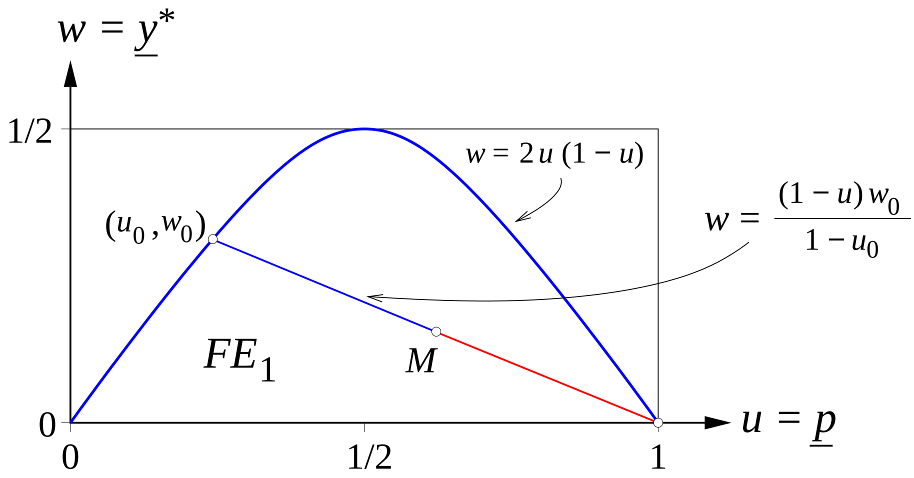

Figure 2 The feasible set FE_1.

consists of attainable pairs contained in FE_1. Since any point (u,w) in FE_1 lies on such a line segment, we conclude that *every point in* FE_1 *is attainable*. We shall refer to FE_1 as *the feasible set*, which is shown in Figure 2.

We now attempt to characterize the points of the feasible set according to whether the CJT is satisfied or not. For that we first define:

Definition 22.

- *The* strong CJT *set, denoted by* $sCJT$*, is the set of all points* $(u,w) \in FE_1$ *such that any pair* (X,P) *with parameters* $\underline{p} = u$ *and* $\underline{y}^* = w$ *satisfies the* CJT.
- *The* weak CJT *set, denoted by* $wCJT$*, is the set of all points* $(u,w) \in FE_1$ *for which there exists a pair* (X,P) *with parameters* $\underline{p} = u$ *and* $\underline{y}^* = w$ *that satisfies the* CJT.

We denote $-sCJT = FE_1 \backslash sCJT$ *and* $-wCJT = FE_1 \backslash wCJT$.

For example, as we shall see later (see Proposition 24), $(1,0) \in sCJT$ and $(1/2,0) \in wCJT$.

By Theorem 9, if $u < 1/2 + 1/2w$, then $(u,w) \in -wCJT$. Next we observe that if (u_0,w_0) is on the parabola $w = 2u(1-u)$ and M is the midpoint of the segment $[(u_0,w_0),(1,0)]$, then by the proof of Theorem 20, the segment $[M,(1,0)] \subseteq wCJT$ (see Figure 2). To find the upper boundary of the union of all these segments, that is, the locus of the mid points M in Figure 2, we eliminate (u_0,w_0) from the equations $w_0 = 2u_0(1-u_0)$, and $(u,w) = 1/2(u_0,w_0) + 1/2(1,0)$, and obtain

$$w = 2(2u-1)(1-u) \tag{26}$$

This is a parabola with maximum $1/4$ at $u = 3/4$. The slope of the tangent at $u = 1/2$ is 2; that is, the tangent of the parabola at that point is the line $w = 2u - 1$ defining the region $-wCJT$. Finally, a careful examination of the proof of Theorem 20, reveals that for every (u_0, w_0) on the parabola $w = 2u(1-u)$, the line segment $[(u_0, w_0), M]$ is in $-sCJT$ (see Figure 2).

Our analysis so far leads to the conclusions summarized in Figure 3 describing the feasibility and and regions of CJT possibility for all pairs (X, P).

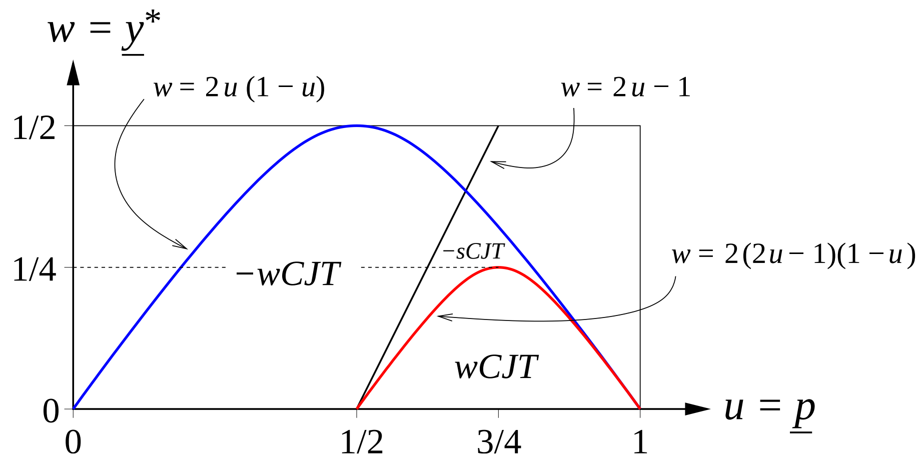

Figure 3 Regions of possibility of CJT in L_1.

Figure 3 is not complete in the sense that the regions $wCJT$ and $-sCJT$ are not disjoint, as it may mistakenly appear in the figure. More precisely, we complete Definition 22 by defining:

Definition 23. *The* mixed *CJT set, denoted by $mCJT$, is the set of all points $(u, w) \in FE_1$ for which there exists a pair (X, P) with parameters $\underline{p} = u$ and $\underline{y}^* = w$ that satisfies the CJT, and a pair $(\hat{X}, \hat{P})$ with parameters $\underline{\hat{p}} = u$ and $\underline{\hat{y}}^* = w$ for which the CJT is violated.*

Then the regions $sCJT$, $-wCJT$, and $mCJT$ are disjoint and form a partition of the feasible set of all distributions FE_1

$$FE_1 = -wCJT \cup sCJT \cup mCJT \tag{27}$$

To complete the characterization we have to find the regions of this partition, and for that it remains to identify the region $mCJT$ since, by definition, $wCJT \backslash mCJT \subset sCJT$ and $-sCJT \backslash mCJT \subset -wCJT$.

Proposition 24. *All three regions $sCJT$, $-wCJT$, and $mCJT$ are not empty.*

Proof. As can be seen from Figure 3, the region $-wCJT$ is clearly not empty; it contains for example the points $(0,0)$ and $(1/2, 1/2)$. As we remarked already (following Definition 22), the region $sCJT$ contains the point $(1,0)$. This point corresponds to a unique pair (X, P), in which $X_i = 1$ for all i with probability 1, that trivially satisfies the CJT). The region $mCJT$ contains the point $(1/2,0)$. To see this recall that the Berend and Paroush's necessary and sufficient condition for CJT in the independent case (see Remark 5) is

$$\lim_{n\to\infty}\sqrt{n}(\overline{p}_n-\frac{1}{2})=\infty \tag{28}$$

First consider the pair $(\tilde{X},\tilde{P})$ in which $(\tilde{X}_i)_{i=1}^{\infty}$ are $i.i.d$ with $P(\tilde{X}_i=1)=1/2$ and $P(\tilde{X}_i=0)=1/2$. Clearly $\sqrt{n}(\overline{p}_n-\frac{1}{2})=0$ for all n and hence condition (28) is not satisfied, implying that the CJT is not satisfied.

Now consider (X,P) in which $X=(1,1,0,1,0,1\cdots)$ with probability 1. This pair corresponds to the point $(1/2,0)$ since

$$\overline{X}_n=\overline{p}_n=\begin{cases}\frac{1}{2}+\frac{1}{n} & \text{if } n \text{ is even}\\ \frac{1}{2}+\frac{1}{2n} & \text{if } n \text{ is odd}\end{cases},$$

and hence $\underline{p}=1/2$ and $\underline{y}^*=0$. Finally this sequence satisfies the CJT as $\overline{X}_n>\frac{1}{2}$ with probability one for all n. □

6.2 Feasibility and characterization in L_2

Replacing $\underline{y}^*=\liminf_{n\to\infty}E|\overline{X}_n-\overline{p}_n|$ by the parameter $\underline{y}=\liminf_{n\to\infty}E(\overline{X}_n-\overline{p}_n)^2$, we obtain results in the space of points $(\underline{p},\underline{y})$ similar to those obtained in the previous section in the space $(\underline{p},\underline{y}^*)$.

Given a sequence of binary random variables X with its joint distribution P, we first observe that for any $i\neq j$,

$$Cov(X_i,X_j)=E(X_iX_j)-p_ip_j\leq\min(p_i,p_j)-p_ip_j.$$

Therefore,

$$E(\overline{X}_n-\overline{p}_n)^2 = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j\neq i}Cov(X_i,X_j)+\sum_{i=1}^{n}p_i(1-p_i)\right\} \tag{29}$$

$$\leq \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j\neq i}[\min(p_i,p_j)-p_ip_j]+\sum_{i=1}^{n}p_i(1-p_i)\right\}. \tag{30}$$

We claim that the maximum of the last expression (30), under the condition $\sum_{i=1}^{n}p_i=\overline{p}_n$, is $\overline{p}_n(1-\overline{p}_n)$. This is attained when $p_1=\cdots=p_n=\overline{p}_n$. To see that this is indeed the maximum, assume to the contrary that the maximum is attained at $\tilde{p}=(\tilde{p}_1,\cdots,\tilde{p}_n)$ with $\tilde{p}_i\neq\tilde{p}_j$ for some i and j. Without loss of generality assume that: $\tilde{p}_1\leq\tilde{p}_2\leq\cdots\leq\tilde{p}_n$ with $\tilde{p}_1<\tilde{p}_j$ and $\tilde{p}_1=\tilde{p}_\ell$ for $\ell<j$. Let $0<\varepsilon<(\tilde{p}_j-\tilde{p}_1)/2$ and define $p^*=(p_1^*,\cdots,p_n^*)$ by $p_1^*=\tilde{p}_1+\varepsilon,\ \ p_j^*=\tilde{p}_j-\varepsilon$, and $p_\ell^*=\tilde{p}_\ell$ for $\ell\notin\{1,j\}$. A tedious, but straightforward, computation shows that the expression (30) is higher for p^* than for $\tilde{p}$, in contradiction to the assumption that it is maximized at $\tilde{p}$. We conclude that

$$E(\overline{X}_n-\overline{p}_n)^2\leq\overline{p}_n(1-\overline{p}_n).$$

Let now $(\overline{p}_{n_k})_{k=1}^{\infty}$ be a subsequence converging to $\underline{p}$; then

$$\begin{aligned}\underline{y} &= \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 \leq \liminf_{k\to\infty} E(\overline{X}_{n_k} - \overline{p}_{n_k})^2 \\ &\leq \liminf_{k\to\infty} \overline{p}_{n_k}(1-\overline{p}_{n_k}) = \underline{p}(1-\underline{p}).\end{aligned}$$

We state this as a theorem:

Theorem 25. *For every pair (X,P), the corresponding parameters $(\underline{p},\underline{y})$ satisfy $\underline{y} \leq \underline{p}(1-\underline{p})$.*

Next we have the analogue of Theorem 20, proved in the same way.

Theorem 26. *Let $t \in [0,\frac{1}{2}]$. If F is a distribution with parameters $(\underline{p},\underline{y})$, then there exists a distribution H with parameters $\underline{\tilde{p}} = 1-t+t\underline{p}$ and $\underline{\tilde{y}} = t^2\underline{y}$ that satisfy the CJT.*

We can now construct Figure 4, which is the analogue of Figure 2 in the L_2 space $(\underline{p},\underline{y})$. The feasible set in this space is

$$FE_2 = \{(u,w) | 0 \leq u \leq 1, \ \ 0 \leq w \leq u(1-u)\} \tag{31}$$

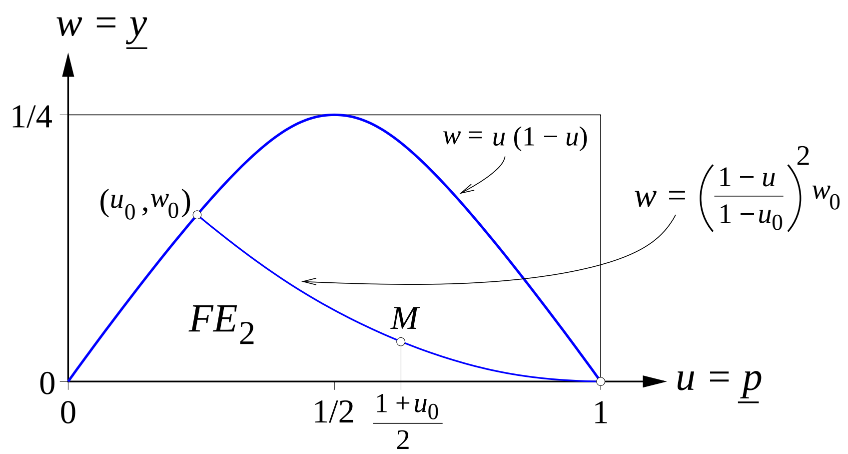

Figure 4 The feasible set FE_2.

The geometric locus of the midpoints M in Figure 4 is derived from (1) $u = \frac{1}{2}u_0 + \frac{1}{2}$, (2) $w = \frac{1}{4}w_0$, and (3) $w_0 = u_0(1-u_0)$ and is given by $w = \frac{1}{2}(2u-1)(1-u)$. This yields Figure 5, which is the analogue of Figure 3. Note, however, that unlike in Figure 3, the straight line $w = u - \frac{1}{2}$ *is not tangential* to the small parabola $w = (u-\frac{1}{2})(1-u)$ at $(\frac{1}{2},0)$.

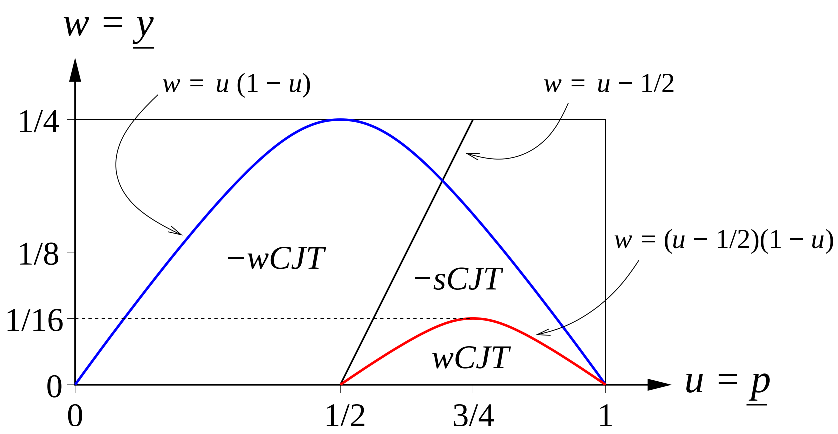

Figure 5 Regions of possibility of CJT in L_2.

The next step toward determining the region $mCJT$ in the L_2 space (Figure 5) is the following:

Proposition 27. *For any* $(u,w) \in \{(u,w)|0 < u < 1\,;\ 0 \le w \le u(1-u)\}$, *there is a pair* (Z,H) *such that:*

(i) $E(Z_i) = u, \forall i$.

(ii) $\liminf_{n\to\infty} E(\overline{Z}_n - u)^2 = w$.

(iii) *The distribution* H *does not satisfy the* CJT.

Proof. For $0 < u < 1$,

- let (X, F_0) be given by $X_1 = X_2 = \ldots = X_n = \ldots$ and $E(X_i) = u$;
- let (Y, F_1) be a sequence of of $i.i.d.$ random variables $(Y_i)_{i=1}^{\infty}$ with expectation u.
- For $0 < t \le 1$ let (Z^t, H^t) be the pair in which $Z_i^t = tX_i + (1-t)Y_i$ for $i = 1, 2, \ldots$ and H^t is the product distribution $H^t = F_0 \times F_1$ (that is, the X and the Y sequences are independent).

Note first that $E(Z_i^t) = u$ for all i and

$$\lim_{n\to\infty} E(\overline{Z^t}_n - u)^2 = \lim_{n\to\infty}\left((1-t)\frac{u(1-u)}{n} + tu(1-u)\right) = tu(1-u),$$

and therefore the pair (Z^t, H^t) corresponds to the point (u,w) in the L_2 space, where $w = tu(1-u)$ ranges in $(0, u(1-u))$ as $0 < t \le 1$.

Finally, (Z^t, H^t) does not satisfy the CJT since for all n,

$$Pr(\overline{Z^t}_n > \frac{1}{2}) \le 1 - Pr(Z_1^t = Z_2^t = \ldots = 0) = 1 - t(1-u) < 1.$$

As this argument does not apply for $t = 0$ it remains to prove that, except for $(1,0)$, to any point $(u,0)$ on the x axis corresponds a distribution that does not satisfy the CJT. For $0 \leq u \leq 1/2$, the sequence (Y, F_1) of of $i.i.d.$ random variables $(Y_i)_{i=1}^{\infty}$ with expectation u does not satisfy the CJT, as follows from the result of Berend and Paroush (1998). For $1/2 < u < 1$ such a sequence of $i.i.d.$ random variables does satisfy the CJT and we need the following more subtle construction:

Given the two sequences (X, F_0) and (Y, F_1) defined above we construct a sequence $Z = (Z_i)_{i=1}^{\infty}$ consisting of alternating blocks of X_i-s and Y_i-s, with the probability distribution on Z being that induced by the product probability $H = F_0 \times F_1$. Clearly $E(Z_i) = u$ for all i, in particular $\overline{p}_n = u$ for all n and $\underline{p} = u$. We denote by B_ℓ the set of indices of the ℓ-th block and its cardinality by b_ℓ. Thus $n(\ell) = \Sigma_{j=1}^{\ell} b_j$ is the index of Z_i at the end of the ℓ-th block. Therefore

$$B_{\ell+1} = \{n(\ell)+1, \ldots, n(\ell)+b_{\ell+1})\} \quad \text{and} \quad n(\ell+1) = n(\ell)+b_{\ell+1}.$$

Define the block size b_ℓ inductively by:

1. $b_1 = 1$, and for $k = 1, 2, \ldots$,
2. $b_{2k} = k\Sigma_{j=1}^{k} b_{2j-1}$ and $b_{2k+1} = b_{2k}$.

Finally we define the sequence $Z = (Z_i)_{i=1}^{\infty}$ to consist of X_i-s in the odd blocks and Y_i-s in the even blocks, that is,

$$Z_i = \begin{cases} X_i & \text{if} \quad i \in B_{2k-1} \quad \text{for some} \;\; k = 1, 2, \ldots \\ Y_i & \text{if} \quad i \in B_{2k} \quad\;\; \text{for some} \;\; k = 1, 2, \ldots \end{cases}$$

Denote by $n_x(\ell)$ and $n_y(\ell)$ the number of X coordinates and Y coordinates respectively in the sequence Z at the end of the ℓ-th block and by $n(\ell) = n_x(\ell) + n_y(\ell)$ the number of coordinates at the end of the ℓ-th block of Z. It follows from 1 and 2 (in the definition of b_ℓ) that for $k = 1, 2, \ldots$,

$$n_x(2k-1) = n_y(2k-1) + 1 \tag{32}$$

$$\frac{n_x(2k)}{n_y(2k)} \leq \frac{1}{k} \quad \text{and hence also} \quad \frac{n_x(2k)}{n(2k)} \leq \frac{1}{k} \tag{33}$$

It follows from (32) that at the end of each odd-numbered block $2k-1$, there is a majority of X_i coordinates that with probability $(1-u)$ will all have the value 0. Therefore,

$$Pr\left(\overline{Z}_{n(2k-1)} < \frac{1}{2}\right) \geq (1-u) \quad \text{for} \;\; k = 1, 2, \ldots,$$

and hence

$$\liminf_{n \to \infty} Pr\left(\overline{Z}_n > \frac{1}{2}\right) \leq u < 1;$$

that is, (Z,H) does not satisfy the CJT.
It remains to show that

$$\underline{y} = \liminf_{n\to\infty} E(\overline{Z}_n - \overline{p}_n)^2 = 0.$$

To do so, we show that the subsequence of $\left\{E((\overline{Z}_n - \overline{p}_n)^2)\right\}_{n=1}^{\infty}$ corresponding to the end of the even-numbered blocks converges to 0, namely,

$$\lim_{k\to\infty} E(\overline{Z}_{n(2k)} - \overline{p}_{n(2k)})^2 = 0.$$

Indeed,

$$E(\overline{Z}_{n(2k)} - \overline{p}_{n(2k)})^2 = E\left(\frac{n_x(2k)}{n(2k)}(X_1 - u) + \frac{1}{n(2k)}\Sigma_{i=1}^{n_y(2k)}(Y_i - u)\right)^2.$$

Since the Y_i-s are $i.i.d.$ and independent of X_1 we have

$$E(\overline{Z}_{n(2k)} - \overline{p}_{n(2k)})^2 = \frac{n_x^2(2k)}{n^2(2k)}u(1-u) + \frac{n_y(2k)}{n^2(2k)}u(1-u),$$

and by property (33) we get finally:

$$\lim_{k\to\infty} E(\overline{Z}_{n(2k)} - \overline{p}_{n(2k)})^2 \leq \lim_{k\to\infty}\left(\frac{1}{k^2}u(1-u) + \frac{1}{n(2k)}u(1-u)\right) = 0,$$

concluding the proof of the proposition. □

Proposition 27 asserts that for every point (u,w) in Figure 5, except for the point $(1,0)$, there is a distribution with these parameters that does not satisfy the CJT. This was known from our previous results in the regions $-wCJT$ and $-sCJT$. As for the region $wCJT$, Proposition 27 and Theorem 26 yield the following conclusions, presented in Figure 6.

Corollary 28. *1. The region below the small parabola in Figure 5, with the exception of the point $(1,0)$, is in $mCJT$, that is,*

$$\left\{(\underline{p},\underline{y})|\frac{1}{2} \leq \underline{p} < 1;\ \text{ and }\ \underline{y} \leq \frac{1}{2}(2\underline{p}-1)(1-\underline{p})\right\} \subseteq mCJT.$$

2. The point $(\underline{p},\underline{y}) = (1,0)$ is the only *point in $sCJT$. It corresponds to a single sequence with $X_1 = \cdots = X_n = \cdots$ with $F(X_i = 1) = 1$.*

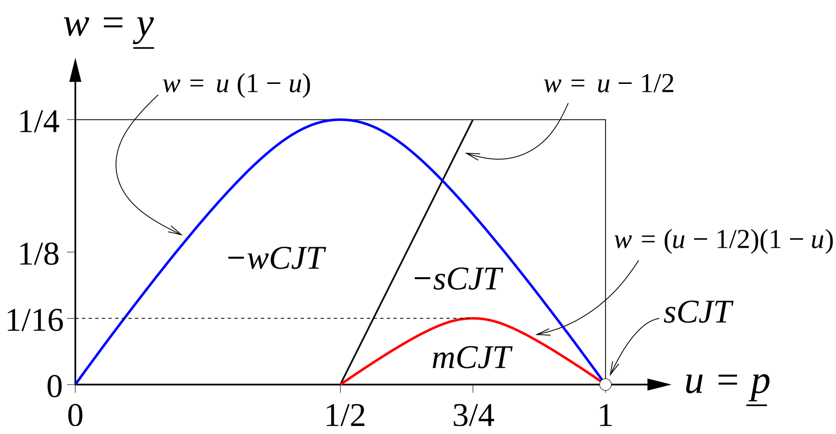

Figure 6 $mCJT$ and $sCJT$ in the L_2 space.

7 Exchangeable variables

In this section we fully characterize the distributions of sequences $X = (X_1, X_2, ..., X_n, ...)$ of *exchangeable* variables that satisfy the CJT. We first recall:

Definition 29. *A sequence of random variables $X = (X_1, X_2, ..., X_n, ...)$ is* exchangeable *if for every n and every permutation $(k_1, \ldots, k_n)$ of $(1, \ldots, n)$, the finite sequence $(X_{k_1}, \ldots, X_{k_n})$ has same n-dimensional probability distribution as $(X_1, \ldots, X_n)$.*

We shall make use of the following characterization theorem due to de Finetti (see, e.g., Feller (1966), Vol. II, page 225).

Theorem 30. *A sequence of binary random variables $X = (X_1, X_2, ..., X_n, ...)$ is exchangeable if and only if there is a probability distribution F on $[0, 1]$ such that for every n:*

$$Pr(X_1 = \cdots = X_k = 1,\ X_{k+1} = \ldots = X_n = 0) = \int_0^1 \theta^k (1-\theta)^{n-k} dF \tag{34}$$

$$Pr(X_1 + \cdots + X_n = k) = \binom{n}{k} \int_0^1 \theta^k (1-\theta)^{n-k} dF \tag{35}$$

The usefulness of de Finetti for our purposes is that it enables an easy projection of the distribution into our L_2 space:

Theorem 31. *Let $X = (X_1, X_2, ..., X_n, ...)$ be a sequence of exchangeable binary random variables and let F be the corresponding distribution function in de Finetti's theorem. Then,*

$$\underline{y} = \lim_{n\to\infty} E(\overline{X}_n - u)^2 = V(F), \tag{36}$$

where

$$u = \int_0^1 \theta dF \quad \text{and} \quad V(F) = \int_0^1 (\theta - u)^2 dF.$$

Proof. We have

$$u = E(X_i) = Pr(X_i = 1) = \int_0^1 x dF \ ; \ V(X_i) = u(1-u)$$

and for $i \neq j$,

$$Cov(X_i, X_j) == Pr(X_i = X_j = 1) - u^2 = \int_0^1 x^2 dF - u^2 = V(F).$$

So,

$$\begin{aligned} E(\overline{X}_n - u)^2 &= E\left(\frac{1}{n}\Sigma_1^n (X_i - u)\right)^2 \\ &= \frac{1}{n^2}\Sigma_1^n V(X_i) + \frac{1}{n^2}\Sigma_{i \neq j} Cov(X_i, X_j) \\ &= \frac{nu(1-u)}{n^2} + \frac{n(n-1)}{n^2} V(F), \end{aligned}$$

which implies equation (36). □

We can now state the characterization theorem:

Theorem 32. *A sequence $X = (X_1, X_2, ..., X_n, ...)$ of binary exchangeable random variables with a corresponding distribution $F(\theta)$ satisfies the CJT if and only if*

$$Pr\left(\frac{1}{2} < \theta \leq 1\right) = 1, \tag{37}$$

that is, if and only if the support of F is in the open semi-interval $(1/2, 1]$.

Proof. The "only if" part follows from the fact that any sequence $X = (X_1, X_2, ..., X_n, ...)$ of binary $i.i.d.$ random variables with expectation $E(X_i) = \theta \leq 1/2$, violates the CJT (by the Berend and Paroush's necessary condition).

To prove that a sequence satisfying condition (37) also satisfies the CJT, note that for $0 < \varepsilon < 1/2$,

$$Pr\left(\overline{X}_n > \frac{1}{2}\right) \geq Pr\left(\theta \geq \frac{1}{2} + \varepsilon\right) Pr\left(\overline{X}_n > \frac{1}{2} | \theta \geq \frac{1}{2} + \varepsilon\right). \tag{38}$$

For the second term in (38) we have:

$$\begin{aligned} Pr\left(\overline{X}_n > \frac{1}{2}|\theta \geq \frac{1}{2}+\varepsilon\right) &= \Sigma_{k>\frac{n}{2}} Pr\left(X_1+\cdots+X_k = k|\theta \geq \frac{1}{2}+\varepsilon\right) & (39)\\ &= \Sigma_{k>\frac{n}{2}} \binom{n}{k} \int_{\frac{1}{2}+\varepsilon}^{1} \theta^k(1-\theta)^{n-k} dF & (40)\\ &= \int_{\frac{1}{2}+\varepsilon}^{1} \left[\Sigma_{k>\frac{n}{2}} \binom{n}{k} \theta^k(1-\theta)^{n-k}\right] dF & (41)\\ &:= \int_{\frac{1}{2}+\varepsilon}^{1} S_n(\theta) dF & (42) \end{aligned}$$

Now, using Chebyshev's inequality we have:

$$\begin{aligned} S_n(\theta) = Pr\left(\overline{X}_n > \frac{1}{2}|\theta\right) &\geq Pr\left(\overline{X}_n > \frac{1}{2}+\varepsilon|\theta\right) & (43)\\ &\geq 1-\frac{V(\overline{X}_n|\theta)}{(\theta-\frac{1}{2}-\varepsilon)^2} = 1-\frac{\theta(1-\theta)}{n(\theta-\frac{1}{2}-\varepsilon)^2} & (44) \end{aligned}$$

Since the last expression in (44) converges to 1 uniformly on $[1/2+\varepsilon,1]$ as $n \to \infty$, taking the limit $n \to \infty$ of (42) and using (44) we have:

$$\lim_{n\to\infty} Pr\left(\overline{X}_n > \frac{1}{2}|\theta \geq \frac{1}{2}+\varepsilon\right) \geq \int_{\frac{1}{2}+\varepsilon}^{1} dF = Pr\left(\theta \geq \frac{1}{2}+\varepsilon\right). \qquad (45)$$

From (38) and (45) we have that for and fixed $\varepsilon > 0$,

$$\lim_{n\to\infty} Pr\left(\overline{X}_n > \frac{1}{2}\right) \geq \left[Pr\left(\theta \geq \frac{1}{2}+\varepsilon\right)\right]^2. \qquad (46)$$

Since (46) must hold for all $1/2 > \varepsilon > 0$, and since $Pr\left(\frac{1}{2} < \theta \leq 1\right) = 1$, we conclude that

$$\lim_{n\to\infty} Pr\left(\overline{X}_n > \frac{1}{2}\right) = 1, \qquad (47)$$

i.e., the sequence $X = (X_1, X_2, ..., X_n, ...)$ satisfies the CJT. □

To draw the consequences of Theorem 32 we prove first the following:

Proposition 33. *Any distribution F of a variable θ in $[1/2,1]$ satisfies*

$$V(F) \leq (u-\frac{1}{2})(1-u), \qquad (48)$$

where $u = E(F)$, and equality holds in (48) only for F in which

$$Pr(\theta = \frac{1}{2}) = 2(1-u) \quad \text{and} \quad Pr(\theta = 1) = 2u-1. \qquad (49)$$

Proof. We want to show that

$$\int_{1/2}^{1} \theta^2 dF(\theta) - u^2 \leq (u - \frac{1}{2})(1-u), \tag{50}$$

or, equivalently,

$$\int_{1/2}^{1} \theta^2 dF(\theta) - \frac{3}{2}u + \frac{1}{2} \leq 0. \tag{51}$$

Replacing $u = \int_{1/2}^{1} \theta dF(\theta)$ and $\frac{1}{2} = \int_{1/2}^{1} \frac{1}{2} dF(\theta)$, inequality (50) is equivalent to

$$\int_{1/2}^{1} (\theta^2 - \frac{3}{2}\theta + \frac{1}{2}) dF(\theta) := \int_{1/2}^{1} g(\theta) dF(\theta) \leq 0. \tag{52}$$

The parabola $g(\theta)$ is convex and satisfies $g(1/2) = g(1) = 0$ and $g(\theta) < 0$ for all $1/2 < \theta < 1$, which proves (52). Furthermore, equality to 0 in (52) is obtained only when F is such that $Pr(1/2 < \theta < 1) = 0$, and combined with $u = E(F)$ this implies (49). □

The next proposition provides a sort of an inverse to proposition 33.

Proposition 34. *For any pair* (u, w) *where* $1/2 < u \leq 1$ *and* $0 \leq w < (u - 1/2)(1-u)$, *there is a distribution* $F(\theta)$ *on* $(1/2, 1]$ *such that* $E(F) = u$ *and* $V(F) = w$.

Proof. For $u = 1$, the only point in this region is when $w = 0$ and for this point $(1, 0)$ the claim is trivially true (with the distribution $Pr(\theta = 1) = 1$), and so it suffices to consider only the case $u < 1$. Given (u, w), for any y satisfying $1/2 < y \leq u < 1$ define the distribution F_y for which

$$Pr(\theta = y) = (1-u)/(1-y) \quad \text{and} \quad Pr(\theta = 1) = (u-y)/(1-y).$$

This distribution satisfies $E(F_y) = u$ and it remains to show that we can choose y so that $V(F_y) = w$. Indeed,

$$V(F_y) = \frac{1-u}{1-y} y^2 + \frac{u-y}{1-y} - u^2.$$

For a given $u < 1$ this is a continuous function of y satisfying: $\lim_{y \to u} V(F_y) = 0$ and $\lim_{y \to 1/2} V(F_y) = (u - 1/2)(1-u)$. Therefore, for $0 \leq w < (u - 1/2)(1-u)$, there is a value y^* for which $V(F_{y^*}) = w$. □

The geometric expression of Theorem 32 can now be stated as follows:
In the L_2 plane of $(\underline{p}, \underline{y})$ let

$$A = \left\{ (\underline{p}, \underline{y}) | \frac{1}{2} < \underline{p} \leq 1; \text{ and } \underline{y} < (\underline{p} - \frac{1}{2})(1 - \underline{p}) \right\} \bigcup \{(1,0)\}$$

This is the region strictly below the small parabola in Figure 6, excluding $(1/2, 0)$ and adding $(1, 0)$.

Theorem 35. *1. Any exchangeable sequence of binary random variables that satisfy the CJT corresponds to $(\underline{p}, \underline{y}) \in A$.*

2. To any $(\underline{p}, \underline{y}) \in A$ there exists an exchangeable sequence of binary random variables with parameters $(\underline{p}, \underline{y})$ that satisfy the CJT.

Proof. The statements of the theorems are trivially true for the point $(1,0)$, as it corresponds to the unique distribution: $Pr(X_1 = \ldots = X_n \ldots) = 1$, which is both exchangeable and satisfies the CJT. For all other points in A,

- Part 1. follows de Finetti's Theorem 30, Theorem 32 and Proposition 33.
- Part 2. follows de Finetti's Theorem 30, Theorem 32 and Proposition 34.

□

Remark 36. *Note that Theorem 26 and part (ii) of Theorem 35 each establish the existence of distributions satisfying the CJT in the interior of the region below the small parabola in Figure 6. The two theorems exhibit examples of such a distribution: while the proof of part (ii) of Theorem (35) provides to each point $(\underline{p}, \underline{y})$ in this region a corresponding exchangeable distribution that satisfies the CJT, Theorem (26), provides other distributions satisfying the CJT that are clearly not exchangeable.*

8 General interlacing

We now generalize the main construction of the proof of Theorem 20. This may be useful in advancing our investigations.

Definition 37. *Let $X = (X_1, X_2, ..., X_n, ...)$ be a sequence of binary random variables with joint probability distribution F and let $Y = (Y_1, Y_2, ..., Y_n, ...)$ be another sequence of binary random variables with joint distribution G. For $t \in [0,1]$, the t-*interlacing *of (X,F) and (Y,G) is the pair $(Z,H) := (X,F) *_t (Y,G)$ where for $n = 1, 2, \ldots$,*

$$Z_n = \begin{cases} X_{\lfloor tn \rfloor} & \text{if} \quad \lfloor tn \rfloor > \lfloor t(n-1) \rfloor \\ Y_{\lceil (1-t)n \rceil} & \text{if} \quad \lceil (1-t)n \rceil > \lceil (1-t)(n-1) \rceil \end{cases}, \tag{53}$$

and $H = F \times G$ is the product probability distribution of F and G.

The following lemma is a direct consequence of Definition 37.

Lemma 38. *If (X,F) and (Y,G) satisfy the CJT then for any $t \in [0,1]$ the pair $(Z,H) = (X,F) *_t (Y,G)$ also satisfies the CJT.*

Proof. We may assume that $t \in (0,1)$. Note that

$$\left\{\omega|\overline{Z}_n(\omega) > \frac{1}{2}\right\} \supseteq \left\{\omega|\overline{X}_{\lfloor tn \rfloor}(\omega) > \frac{1}{2}\right\} \cap \left\{\omega|\overline{Y}_{\lceil (1-t)n \rceil}(\omega) > \frac{1}{2}\right\}$$

By our construction and the fact that both (X,F) and (Y,G) satisfy the CJT,

$$\lim_{n\to\infty} F\left(\overline{X}_{\lfloor tn \rfloor} > \frac{1}{2}\right) = 1 \text{ and } \lim_{n\to\infty} G\left(\overline{Y}_{\lceil (1-t)n \rceil} > \frac{1}{2}\right) = 1.$$

As

$$H\left(\overline{Z}_n > \frac{1}{2}\right) \geq F\left(\overline{X}_{\lfloor tn \rfloor} > \frac{1}{2}\right) \cdot G\left(\overline{Y}_{\lceil (1-t)n \rceil} > \frac{1}{2}\right),$$

the proof follows. ☐

Corollary 39. *The region wCJT is* star-convex *in the L_1 space . Hence, in particular, it is path-connected in this space.*

Proof. Let (u,w) be a point in *wCJT* in the L_1 space. Then, there exists a pair (X,F) which satisfies CJT, where X is the sequence of binary random variables with joint probability distribution F satisfying $\underline{p} = u$ and $\underline{y}^* = w$. By Remark 21, Lemma 38, and the proof of Theorem 20, the line segment $[(u,w),(1,0)]$ is contained in *wCJT*, proving that *wCJT* is star-convex. ☐

Corollary 40. *The region wCJT is* path-connected *in the L_2 space.*

Proof. In the L_2 space a point (u,w) corresponds to $\underline{p} = u$ and $\underline{y} = w$. By the same arguments as before, the arc of the parabola $w = ((1-u)/(1-u_0))^2 w_0$ connecting (u,w) to $(1,0)$ (see Figure 4) is contained in *wCJT*, and thus *wCJT* is path-connected. ☐

9 CJT in the space of all probability distributions.

We look at the space $S_p = \{0,1\}^\infty$ already introduced in the proof of Proposition 27 on page 17. We consider S_p both as a measurable product space and as a topological product space. Let $\mathscr{P}$ be the space of all probability distributions on S_p. $\mathscr{P}$ is a compact metric space in the weak topology. Let $\mathscr{P}_1 \subseteq \mathscr{P}$ be the set of all probability distributions in $\mathscr{P}$ that satisfy the CJT and let $\mathscr{P}_2 = \mathscr{P} \setminus \mathscr{P}_1$ be the set of all probability distributions in $\mathscr{P}$ that do not satisfy the CJT.

Lemma 41. *If P_1 and P_2 are two distributions in $\mathscr{P}$, and if $P_2 \in \mathscr{P}_2$ then for any $0 < t < 1$, the distribution $P_3 = tP_1 + (1-t)P_2$ is also in $\mathscr{P}_2$.*

Proof. For $n = 1, 2, \cdots$, let

$$B_n = \left\{ x = (x_1, x_2, \cdots) \in S_p | \frac{1}{n} \sum_{i=1}^{n} x_i > \frac{1}{2} \right\} \tag{54}$$

Since $P_2 \in \mathscr{P}_2$, there exists a subsequence $(B_{n_k})_{k=1}^{\infty}$ and $\varepsilon > 0$ such that $P_2(B_{n_k}) \leq 1 - \varepsilon$ for $k = 1, 2, \cdots$. Then

$$P_3(B_{n_k}) = tP_1(B_{n_k}) + (1-t)P_2(B_{n_k}) \leq t + (1-t)(1-\varepsilon) = 1 - \varepsilon(1-t),$$

implying that $P_3 \in \mathscr{P}_2$. □

Corollary 42. *The set $\mathscr{P}_2$ is dense in $\mathscr{P}$ (in the weak topology) and is convex.*

We proceed now to separate $\mathscr{P}_1$ from $\mathscr{P}_2$ in the space $\mathscr{P}$. We first observe that $\mathscr{P}_1$ is convex by its definition and $\mathscr{P}_2$ is convex by Lemma 41. In order to separate $\mathscr{P}_1$ from $\mathscr{P}_2$ by some linear functional we first define the mapping $T : \mathscr{P} \to \mathbf{R}^{\mathscr{N}}$ where $\mathscr{N} = \{1, 2, \ldots\}$ in the following way: For $P \in \mathscr{P}$ let

$$T(P) = \left(\int \chi_{B_1} dP, \int \chi_{B_2} dP, \ldots, \int \chi_{B_n} dP, \ldots \right)$$

where the sets B_n are defined in (54) and χ_{B_n} is the indicator function of the set B_n, that is, for $x \in S_p$ and $n = 1, 2, \ldots$,

$$\chi_{B_n}(x) = \begin{cases} 1 & \text{if } x \in B_n \\ 0 & \text{if } x \notin B_n \end{cases}.$$

The mapping T is affine and continuous when $\mathbf{R}^{\mathscr{N}}$ is endowed with the product topology. Clearly, $T(\mathscr{P}) \subset \ell_\infty$. Also $T(\mathscr{P}_1)$ and $T(\mathscr{P}_2)$ are convex and if $z_1 \in T(\mathscr{P}_1)$ and $z_2 \in T(\mathscr{P}_2)$ then $\liminf_{k\to\infty}(z_{1,k} - z_{2,k}) \geq 0$, where $z_i = (z_{i,1}, z_{i,2} \ldots)$ for $i = 1, 2$.

Let $B : \ell_\infty \to \mathbf{R}$ be any *Banach limit* on ℓ_∞ (see Dunford and Schwartz (1958), p. 73); then B is a continuous linear functional on ℓ_∞ and as $B(z) \geq \liminf_{k\to\infty} z_k$ for every $z \in \ell_\infty$, we have $B(z_1) \geq B(z_2)$ whenever $z_1 \in T(\mathscr{P}_1)$ and $z_2 \in T(\mathscr{P}_2)$. Thus the composition $B \circ T$, which is also an affine function, satisfies

$$P_1 \in \mathscr{P}_1 \quad \text{and} \quad P_2 \in \mathscr{P}_2 \Longrightarrow B \circ T(P_1) \geq B \circ T(P_2).$$

We can now improve upon the foregoing separation result between $T(\mathscr{P}_1)$ and $T(\mathscr{P}_2)$ as follows:

Proposition 43. *For any $y \in T(\mathscr{P}_2)$ there exists $\psi \in \ell_\infty^*$ such that*

(i) $\psi(z_1) \geq \psi(z_2)$ *for every $z_1 \in T(\mathscr{P}_1)$ and $z_2 \in T(\mathscr{P}_2)$;*

(ii) $\psi(z) > \psi(y)$ *for all $z \in T(\mathscr{P}_1)$.*

Proof. Given $y \in T(\mathscr{P}_2)$, let $C = y - T(\mathscr{P}_1)$. Then, C is a convex subset of ℓ_∞ and there exists an $\varepsilon > 0$ such that $\liminf_{k\to\infty} z_k < -\varepsilon$ for every $z \in C$. Let

$$D = \left\{ z \in \ell_\infty | \liminf_{k\to\infty} z_k \geq -\varepsilon \right\};$$

then D is convex and $C \cap D = \phi$. Furthermore, $0 \in D$ is an interior point. Hence by Dunford and Schwartz (1958) p. 417, there exists $\psi \in \ell_\infty^*$, $\psi \neq 0$ such that $\psi(d) \geq \psi(c)$ for all $d \in D$ and $c \in C$. Since $\ell_\infty^+ = \{x \in \ell_\infty | x \geq 0\}$ is a cone contained in D, and ψ is bounded from below on D (and hence on ℓ_∞^+) it follows that $\psi \geq 0$ (that is $\psi(x) \geq 0$ for all $x \geq 0$). Also, as $0 \notin C$ and $0 \in Int(D)$, there exists $\delta > 0$ such that $\psi(d) \geq -\delta$ for all $d \in D$ and $\psi(D) \supseteq (-\delta, \infty)$. Hence $\psi(c) \leq -\delta$ for all $c \in C$, that is, $\psi(y - x) \leq -\delta$ for all $x \in T(\mathscr{P}_1)$.

Finally, the cone $C^* = \{z \in \ell_\infty | \liminf_{k\to\infty} z_k \geq 0\}$ is contained in D; hence ψ is non-negative on C^*. As $T(\mathscr{P}_1) - T(\mathscr{P}_2) \subseteq C^*$, the functional ψ separates weakly between $T(\mathscr{P}_1)$ and $T(\mathscr{P}_2)$. □

Remark 44. *The mapping $\psi \circ T$, which is defined on $\mathscr{P}$, may not be continuous in the weak topology on $\mathscr{P}$. Nevertheless, it may be useful in distinguishing the elements of $\mathscr{P}_1$ from those of $\mathscr{P}_2$.*

References

Austen-Smith, D. and J.S. Banks (1996), "Information aggregation, rationality, and the Condorcet Jury Theorem." The American Political Science Review 90: 34–45.

Berend, D. and J. Paroush (1998), "When is Condorcet Jury Theorem valid ?" Social Choice and Welfare 15: 481–488.

Condorcet, Marquis de (1785), "Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix". De l'Imprimerie Royale, Paris.

Feller, W. (1957, third edition), "An Introduction to Probability Theory and Its Applications" Volume I, John Wiley & Sons, New York.

Feller, W. (1966, second corrected printing), "An Introduction to Probability Theory and Its Applications", Volume II, John Wiley & Sons, New York.

Dunford, N. and J.T. Schwartz (1958), "Linear Operators", Part I, Interscience Publishers, New York.

Ladha, K.K. (1995), "Information pooling through majority rule: Condorcet's Jury Theorem with correlated votes." Journal of Economic Behavior and Organization 26: 353–372.

Laslier, J.-F. and J. Weibull (2008), "Committee decisions: optimality and equilibrium." Mimeo.

Loève, M. (1963, third edition), "Probability Theory", D. Van Norstrand Company, Inc., Princeton, New Jersey.

Myerson, R.B. (1998), "Extended Poisson games and the Condorcet Jury Theorem." Games and Economic Behavior 25: 111–131.

Myerson, R.B. (1998), "Population uncertainty and Poisson games". International Journal of Game Theory 27: 375–392.

Nitzan, S. and J. Paroush (1985), "Collective Decision Making: An Economic Outlook." Cambridge University Press, Cambridge.

Shapley, L.S. and B. Grofman (1984), "Optimizing group judgemental accuracy in the presence of interdependencies." Public Choice 43: 329–343.

Uspensky, J.V. (1937), "Introduction to Mathematical Probability", McGraw-Hill, New York.

Wit, J. (1998), "Rational Choice and the Condorcet Jury Theorem." Games and Economic Behavior 22: 364–376.

Young, P.H. (1997), "Group choice and individual judgements." In 'Perspectives on Public Choice', Mueller, D.C. (ed.), Cambridge University Press, Cambridge, pp. 181-200.

Appendix

9.1 Proof of Proposition 19

Let C be the uniform bound of the variables, i.e., $|X_i| \leq C$ for all i (in our case $C = 1$). Let

$$Q_n = P\left(|\overline{X}_n - \overline{p}_n)| \leq \varepsilon\right) \quad \text{and hence} \quad P\left(|\overline{X}_n - \overline{p}_n)| > \varepsilon\right) = 1 - Q_n.$$

Then

$$E\left(\overline{X}_n - \overline{p}_n\right)^2 \leq C^2(1 - Q_n) + \varepsilon^2 Q_n \leq C^2(1 - Q_n) + \varepsilon^2$$

If LLN holds, then $\lim_{n\to\infty} Q_n = 1$ and thus the limit of the left-hand side is less than or equal to ε. Since this holds for all $\varepsilon > 0$, the limit of the left-hand side is zero. □