# On Bayesian-Nash Equilibria Satisfying the Condorcet Jury Theorem: The Dependent Case

Bezalel Peleg and Shmuel Zamir[1,2]

January 14, 2010

**Abstract**

We investigate sufficient conditions for the existence of Bayesian-Nash equilibria that satisfy the *Condorcet Jury Theorem* ($CJT$). In the Bayesian game $G_n$ among $n$ jurors, we allow for arbitrary distribution on the types of jurors. In particular, any kind of dependency is possible. If each juror $i$ has a "constant strategy", $\sigma^i$ (that is, a strategy that is independent of the size $n \geq i$ of the jury), such that $\sigma = (\sigma^1, \sigma^2, \ldots, \sigma^n \ldots)$ satisfies the $CJT$, then by McLennan (1998) there exists a Bayesian-Nash equilibrium that also satisfies the $CJT$. We translate the $CJT$ condition on sequences of constant strategies into the following problem:

(**) For a given sequence of binary random variables $X = (X^1, X^2, ..., X^n, ...)$ with joint distribution $P$, does the distribution $P$ satisfy the asymptotic part of the $CJT$ ?

We provide sufficient conditions and two general (distinct) necessary conditions for (**). We give a complete solution to this problem when $X$ is a sequence of exchangeable binary random variables.

## 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.*

[1]Center for the Study of Rationality, The Hebrew University of Jerusalem.
[2]We thank Marco Scarsini and Yosi Rinott for drawing our attention to de Finetti's theorem.

We build on some of the literature on this issue in the last thirty years. First we notice that Nitzan and Paroush (1982) and Shapley and Grofman (1984) allow for unequal competencies of the juries. They replace the simple majority committee by weighted majority simple games to maintain the optimality of the voting rule.

Second, we notice the many papers on the dependency among jurors. Among these papers are Shapley and Grofman (1984), Boland, Prochan, and Tong (1989), Ladha (1992, 1993, 1995), Berg (1993a, 1993b), Dietrich and List (2004), Berend and Sapir (2007), and Dietrich (2008). It is widely understood and accepted that the votes of the jurors are often correlated. For example, group deliberation prior to voting is viewed, justifiably, as undermining independence (Grofman, Owen, and Feld (1983), Ladha (1992, 1995), Estlund (1994), and Dietrich and List (2004)). In particular, Dietrich (2008) argues that independence cannot be fully justified in the Condorcet jury model.

Finally, we mention the seminal paper of Austen-Smith and Banks (1996) which incorporated strategic analysis into the Condorcet jury model. This paper had many followers, in particular McLennan (1998), and Duggan and Martinelli (2001) which investigated the Condorcet Jury Theorem ($CJT$) for Bayesian-Nash equilibria (BNE).

In this work, we investigate the $CJT$ for BNE. Unlike Austen-Smith and Banks (1996), we do not assume that the *types* of the voters are independent (given the *state of nature*). Indeed we assume arbitrary dependency among (the types of) jurors. As far as we could ascertain, McLennan (1998) is the only paper that studies the $CJT$ for BNE assuming dependency among the jurors. In fact we rely heavily on McLennan's work; the game among $n$ jurors, is a Bayesian game $G_n$ in which all the players have the same payoff function which is the probability of *correct decision*. Therefore, any $n$-tuple of strategies $\sigma_n = (\sigma_n^1, \ldots, \sigma_n^n)$ that maximizes the common payoff is a BNE (McLennan (1998), Theorem 1). Now consider an infinite sequence of such strategies $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n, \ldots)$ that are BNE for the sequence of games $G_1, G_2, \ldots, G_n, \ldots$ with a jury of growing size. If there exists any other sequence of strategies $\tau = (\tau_1, \tau_2, \ldots, \tau_n, \ldots)$ (not necessarily BNE), that satisfies the $CJT$, then the original sequence $\sigma$ is a sequence (of BNE) that also satisfies the $CJT$. Thus, we may focus on the following problem:

(*) For a given sequence of Bayesian games $G_1, G_2, \ldots, G_n, \ldots$ with an increasing set of jurors, find some sequence of strategies $\tau = (\tau_1, \tau_2, \ldots, \tau_n, \ldots)$ where $\tau_n$ is an $n$-tuple of strategies for the game $G_n$, so that the sequence $(\tau_n)_{n=1}^{\infty}$ satisfies the $CJT$.

In view of the generality and the complexity of our model, we limit ourselves to sequences $\tau$ of "constant" strategies; that is, we assume that $\tau_n^i = \tau_m^i$ if $1 \leq i \leq m \leq n < \infty$. This means that the strategy $\tau_n^i$ of a specific juror $i$ does not change when the size of the jury increases. We shall refer to such a sequence as a "constant sequence." [3] We prove that verifying the $CJT$ for a constant sequence is equivalent to the following problem:

(**) For a given sequence of binary random variables $X = (X^1, X^2, ..., X^n, ...)$ with joint distribution $P$, find whether or not the distribution $P$ satisfies the $CJT$.

[3] The restriction to constant strategies is needed only for the existence results. The sufficient condition as well as the necessary conditions are valid for any infinite sequence of strategies. See Remark 4 on page 6

Note that prior to Austen-Smith and Banks (1996), the analysis of the Condorcet jury problem had focused on problem (**). One general result is that of Berend and Paroush (1998) which characterizes the independent sequences of binary random variables that satisfy the $CJT$.

In this paper we find sufficient conditions for (**). Then we supply two general necessary conditions. However, we do not have a complete characterization of the solution to (**). We do have full characterization (necessary and sufficient conditions) for sequences of *exchangeable* random variables.

Our basic model is introduced in Section 1. The full characterization for the case of exchangeable variables is given in Section 2. In Section 3 we give sufficient conditions for the $CJT$. In Section 4 we develop necessary conditions for the validity of the $CJT$ in two different planes of parameters of the distribution. In Section 5 we prove that these necessary conditions are not sufficient, unless the sequence is of exchangeable random variables. In Section 6 we introduce the notion of *interlacing* of two sequences, which proves to be a s useful tool to construct new classes of distributions that satisfy the $CJT$. In particular we construct rich classes of non-exchangeable sequences that satisfy the $CJT$. Two proofs are given in the Appendix. In the last part of the appendix we clarify the relationship between the $CJT$ and the *Law of Large Numbers* (*LLN*). Basically we show that these are two different properties that do not imply each other in spite of their superficial similarity.

# 1 The basic model

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*: $g$– the defendant is guilty, and $z$– the defendant is innocent. Thus $\Theta = \{g,z\}$ is the set of states of nature. Each juror has two available actions: $c$– to convict the defendant, and $a$– to acquit the defendant; thus $A = \{a,c\}$ is the action set of each of the jurors. Before voting, each jurors gets a private random signal $t^i_j \in T^i = \{t^i_1,\ldots,t^i_{k_i}\}$. In the terminology of games with incomplete information, $T^i$ is the *type set* of juror $i$. The private signals of the jurors may be dependent and may, of course, depend on the state of nature. Again, in the style of games with incomplete information, let $\Omega_n = \Theta \times T^1 \times,\ldots,\times T^n$ be the set of the *states of the world*. That is, a state of the world $\omega = (\theta,t^1,\ldots,t^n)$ consists of the state of nature and the list of types of the $n$ jurors. Let $p^{(n)}$ be the probability distribution (i.e., a common prior) on $\Omega_n$. This is the joint probability distribution on of the state of nature and the signals (types) of all jurors. We assume that the action taken by the finite society of jurors $I = \{1,2,\ldots,n\}$ .i.e., the jury verdict, is determined by the voting rule $V : A^I \to A$, which is the *simple majority* rule (with some tie-breaking procedure such as coin tossing). Finally, to complete the description of the game, we let all jurors have the same payoff function $u : \Theta \times A \to \mathbb{R}$ namely,

$$u(g,c) = u(z,a) = 1 \quad \text{and} \quad u(g,a) = u(z,c) = 0, \quad \forall i \in I$$

This concludes the definition of a game, which we denote by $G_n$. A (pure) strategy of juror $i \in I$ in $G_n$ is a function $s^i : T^i \to A$. We denote by $S^i$ the set of all pure strategies of juror $i \in I$ and by $S = S^1 \times, \ldots, \times S^n$ the set of strategy profiles of the society. The (common) ex-ante payoff for each juror, when the strategy vector $s = (s^1, \ldots, s^n) \in S$ is used, is $E_u = Eu(\theta, V(s^1(t^1), \ldots, s^n(t^n)))$, where $\theta$ is the true state of nature. Note that $E_u$ is precisely the probability correct decision by $I$ when the strategy vector $s$ is used.

**Example 2.** *In the original Condorcet theorem we have $T^i = \{t^i_g, t^i_z\}$; $p^{(n)}(g) = p^{(n)}(z) = 1/2$ and the types are conditionally independent given the state of nature; each has a probability $p > 1/2$ of getting the correct signal. That is,*

$$p^{(n)}(t^i_g|g) = p^{(n)}(t^i_z|z) = p > \frac{1}{2}$$

*Condorcet further assumed that all the jurors vote informatively, that is, use the strategy $s^i(t^i_z) = a$ and $s^i(t^i_g) = c$. In this case, the probability of correct voting, by each juror, is $p$, and as the signals are (conditionally) independent, the $CJT$ follows (for example, by the Law of Large Numbers).*

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: $p^{(2)}(g) = 1/4$ and $p^{(2)}(z) = 3/4$ and then the types of the two jurors are chosen according to a joint probability distribution that depends on the state of nature.

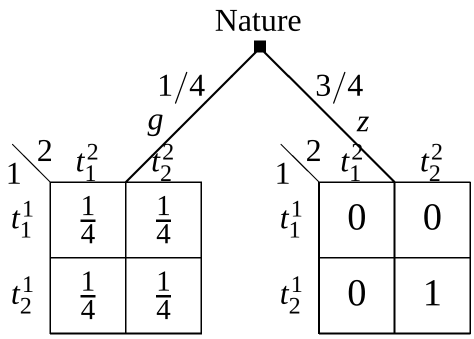


**Figure 1** The probability distribution $p^{(2)}$.

Following the seminal work of Austen-Smith and Banks (1996), we intend to study the $CJT$ via the Bayesian Nash Equilibria (BNE) of the game $G_n$. However, unlike in the case of (conditionally) independent signals, there is no obvious way to find the relevant BNE in the general case of arbitrary dependence. Therefore, our approach will be indirect. Before describing our techniques we first enlarge the set of strategies of the jurors by adding the possibility of mixed strategies. Indeed, it was shown by Wit (1998) that the introduction of mixed strategies may help the realization of the $CJT$.

A mixed strategy[4] for juror $i \in I$, in the game $G_n$, is a function $\sigma_n^i : T^i \to \Delta(A)$, where $\Delta(A)$ is the set of probability distributions on $A$. Denote by $\Sigma_n^i$ the set of all mixed strategies of juror $i$ and by $\Sigma_n = \Sigma_n^1 \times \ldots \times \Sigma_n^n$ the set of mixed strategy vectors (profiles) in the game $G_n$. The (common) ex-ante payoff for each juror, when the strategy vector $\sigma_n = (\sigma_n^1, \ldots, \sigma_n^n) \in \Sigma_n$ is used, is $E_u = Eu(\theta, V(\sigma_n^1(t^1), \ldots, \sigma_n^n(t^n)))$, where $\theta$ is the true state of nature. Again, $E_u$ is precisely the probability of correct decision by $I$ when the strategy vector $\sigma$ is played.

We shall now find a more explicit expression for the payoff $E_u$. Given a strategy vector $\sigma_n = (\sigma_n^1, \ldots, \sigma_n^n) \in \Sigma_n$ we denote by $X_n^i(\sigma_n^i) : \Theta \times T^i \to \{0,1\}$ the indicator of the set of correct voting of juror $i$ when using the mixed strategy $\sigma^i$. That is,

$$X_n^i(\sigma_n^i; \theta_n, t_n^i) = \begin{cases} 1 & \text{if } \theta_n = g \text{ and } \sigma_n^i(t_n^i) = c \text{ or } \theta_n = z \text{ and } \sigma_n^i(t_n^i) = a \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where by a slight abuse of notation we denoted by $\sigma_n^i(t_n^i)$ the realized pure action when juror $i$ of type $t_n^i$ uses mixed strategy $\sigma_n^i$. Given a strategy vector $\sigma_n = (\sigma_n^1, \ldots, \sigma_n^n)$, the probability distribution $p^{(n)}$ on $\Omega_n$ induces a joint probability distribution on the vector of binary random variables $(X_n^1, X_n^2, \ldots, X_n^n)$ which we denote by $p_{\sigma_n}^{(n)}$. Assume now that $n$ is odd; then $E_u$ is given by

$$E_u = p_{\sigma_n}^{(n)}(\Sigma_{i=1}^n X_n^i > \frac{n}{2}).$$

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 $\Omega_{n+1}$ on $\Omega_n$ is the whole $\Omega_n$. 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 $\Omega_n$ 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

$$\Omega = \lim_{\infty \leftarrow n} \Omega_n = \Theta \times T^1 \times \ldots \times T^n \ldots$$

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

Let $(\sigma_n)_{n=1}^\infty$ be an infinite sequence of strategy vectors for an increasing jury. We say that $(\sigma_n)_{n=1}^\infty$ satisfies the (asymptotic part of) $CJT$ if

$$\lim_{n \to \infty} p_{\sigma_n}^{(n)} \left( \Sigma_{i=1}^n X_n^i(\sigma_n^i) > \frac{n}{2} \right) = 1. \tag{2}$$

Our aim in this work is to find sufficient conditions for the existence of a sequence of BNE $(\sigma_n)_{n=1}^\infty$ that satisfy the (asymptotic part of) $CJT$. As far as we know, the only existing result on this general problem is that of Berend and Paroush (1998), which deals only

[4] As a matter of fact, the strategy we define here is a *behavior strategy*, but as the game is clearly a game with *perfect recall*, it follows from Kuhn’s theorem (1953) that any mixed strategy has a payoff equivalent behavior strategy. Thus we (ab)use the term “mixed strategy” which is more familiar in this literature.

with independent jurors. For that, we make use of the following result due to McLennan for games with common interest (which is our case):

**Theorem 3.** *(McLennan (1998)) For $n = 1, 2, \ldots$, if*

$$\sigma_n^* = (\sigma_n^{*1}, \ldots, \sigma_n^{*n}) \in \arg \max_{(\sigma_n^1, \ldots, \sigma_n^n))} E_u(\theta, V(\sigma_n^1(t^1), \ldots, \sigma_n^n(t^n))),$$

*then $\sigma_n^*$ is a Bayesian Nash Equilibrium of the game $G_n$*

This is an immediate application of Theorem 1 in McLennan (1998), which implies that $\sigma_n^*$ is a Nash equilibrium of the type-agent representation of $G_n$. Since by Theorem 3, a Bayesian Nash Equilibrium of $G_n$ maximizes the probability of correct decision, then clearly, if there exists any sequence of strategy vectors $(\sigma_n)_{n=1}^\infty$ that satisfies the asymptotic part of $CJT$, (2), then there is also a sequence $(\sigma_n^*)_{n=1}^\infty$ of BNE that satisfies (2), the asymptotic part of $CJT$.

Our approach in this paper is to provide such a sequence that satisfies the $CJT$. In particular, we shall consider infinite sequences of mixed strategy vectors that are constant with respect to the number of players, that is, $(\sigma_n)_{n=1}^\infty$ such that if $n \geq m$ then $\sigma_n^i = \sigma_m^i$ for all $i \leq m$. Such a constant sequence can be represented as one infinite sequence of strategies $\sigma = (\sigma^1, \sigma^2, \ldots, \sigma^n, \ldots)$, where $\sigma^i$ is the strategy of juror $i$ in all juries that he is a member of (i.e. in all games $G_n$ with $n \geq i$). Whenever we find such a constant sequence that satisfies the $CJT$, it follow, as we argued, that there is a sequence $(\sigma_n^*)_{n=1}^\infty$ of BNE that satisfies (2), the asymptotic part of $CJT$. A constant sequence $(\sigma_n)_{n=1}^\infty$ can be interpreted as a sequence of an increasing jury in which the strategies of the jury members do not change as the jury increases. In addition to their plausibility, we restrict our attention to constant sequences because of the complexity of our model. As we shall see, even with this restriction, we get some interesting results.

**Remark 4.** *As far as we can see, the assumption of constant strategies will be needed only for our existence results (Theorem 8, Corollary 14, and Theorem 27). For the sufficient condition, as well as for the two necessary conditions, we need neither the restriction to constant strategies, nor the assumption on the stationarity of the probabilities $p^{(n)}$ (of $G_n$). The proofs are the same, with the appropriate adjustment of notations; mainly, for a general sequence of strategies $(\sigma_n)_{n=1}^\infty$, the corresponding sequence $X$ of binary random variables, is the sequence of n-vectors of random variables $(X_n^1, \ldots, X_n^n)$ corresponding to the game $G_n$ and the strategy vector $\sigma_n$; that is, $X = (X_1^1\ ;\ X_2^1, X_2^2\ ; \ldots\ ;\ X_n^1, \ldots, X_n^n\ ; \ldots)$. The $CJT$ property is defined, as usual, by equation (2)*

A constant sequence of mixed strategies $\sigma = (\sigma^1, \sigma^2, \ldots, \sigma^n, \ldots)$ naturally yields a sequence of binary random variables $X = (X^1, X^2, \ldots, X^n, \ldots)$ where $X^i := X_n^i(\sigma_n^i; \theta_n, t_n^i)$ is the indicator variable of correct voting of juror $i$ defined in (1, and is independent of $n$ since the strategy is constant). As the $CJT$ is expressed in terms of $X$, we shall mostly be working with this infinite sequence of binary random variables. In fact, working with the infinite sequences $X$ is equivalent to working with the underlying infinite sequences of games and strategy vectors: on the one hand, as we said, a sequence of games $(G_n)_{n=1}^\infty$

and an infinite sequence of constant strategies $\sigma = (\sigma^1, \sigma^2, \ldots, \sigma^n, \ldots)$, yield an infinite sequence $X$ of binary random variables. On the other hand, as we show in Appendix 7.1, for any infinite sequence of binary random variables $X$ there is a sequence of games $(G_n)_{n=1}^{\infty}$ and an infinite sequence of constant strategies $\sigma = (\sigma^1, \sigma^2, \ldots, \sigma^n, \ldots)$ that yield this $X$ as the infinite sequence of the indicators of correct voting.

Let us now briefly remark on the non-asymptotic part of the $CJT$ (see Ben-Yashar and Paroush (2000)). An infinite sequence of mixed strategy vectors $\sigma_n = (\sigma_n^1, \ldots, \sigma_n^n),\ \ n = 1, 2, \ldots$, is said to be *consistent with the majority rule* if for $n = 1, 2, \ldots$,

$$\begin{aligned} p^{(n)}\left(\Sigma_{i=1}^{n} X_n^i(\sigma_n^i) > \frac{n}{2}\right) &> p^{(n)}(X_n^i(\sigma_n^i) = 1); \quad i = 1, \ldots, n \\ p^{(n+1)}\left(\Sigma_{i=1}^{n+1} X_{n+1}^i(\sigma_{n+1}^i) > \frac{n+1}{2}\right) &\geq p^{(n)}\left(\Sigma_{i=1}^{n} X_n^i(\sigma_n^i) > \frac{n}{2}\right); \quad n = 1, 2, \ldots. \end{aligned}$$

In view of the complexity of our model we shall not investigate non-asymptotic consistency with majority rule of infinite sequences of strategies, and shall study only the asymptotic part of the $CJT$.

# 2 Exchangeable variables

In this section we fully characterize the distributions of sequences $X = (X^1, X^2, \ldots, X^n, \ldots)$ of *exchangeable* random binary variables that satisfy the $CJT$. Let us first introduce some notation:
Given a sequence of binary random variables $X = (X^1, X^2, \ldots, X^n, \ldots)$ 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 $\overline{p}_n = (p^1 + p^2, \ldots + p^n)/n$ and $\overline{X}_n = (X^1 + X^2, \ldots + X^n)/n$. Next we recall:

**Definition 5.** *A sequence of random variables $X = (X^1, X^2, \ldots, X^n, \ldots)$ 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 the same n-dimensional probability distribution as $(X^1, \ldots, X^n)$.*

We shall make use of the following characterization theorem due to de Finetti [5] (see, e.g., Feller (1966), Vol. II, page 225).

**Theorem 6.** *A sequence of binary random variables $X = (X^1, X^2, \ldots, X^n, \ldots)$ is exchangeable if and only if there is a probability distribution $F$ on $[0,1]$ such that for every n,*

$$Pr(X^1 = \cdots = X^k = 1,\ X^{k+1} = \ldots = X^n = 0) = \int_0^1 \theta^k (1-\theta)^{n-k} dF \tag{3}$$

$$Pr(X^1 + \cdots + X^n = k) = \binom{n}{k} \int_0^1 \theta^k (1-\theta)^{n-k} dF \tag{4}$$

[5] As far as we know, Ladha (1993) was the first to apply de Finetti's Theorem to exchangeable variables in order to derive (some parts) of $CJT$. However, Ladha investigates only the non-asymptotic part of $CJT$.

Using de Finetti's theorem we can characterize the distributions of sequences of exchangeable binary random variables by their expectation and the asymptotic variance of $\overline{X}_n$.

**Theorem 7.** *Let $X = (X^1, X^2, \ldots, X^n, \ldots)$ be a sequence of exchangeable binary random variables and let $F$ be the corresponding distribution function in de Finetti's theorem. Then,*

$$\underline{y} := \lim_{n\to\infty} E(\overline{X}_n - u)^2 = V(F), \tag{5}$$

*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 (5). □

We can now state the characterization theorem:

**Theorem 8.** *A sequence $X = (X^1, X^2, \ldots, X^n, \ldots)$ of binary exchangeable random variables with a corresponding distribution $F(\theta)$ satisfies the $CJT$ if and only if*

$$Pr\left(\frac{1}{2} < \theta \leq 1\right) = 1, \tag{6}$$

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

*Proof.* The "only if" part follows from the fact that any sequence $X = (X^1, X^2, \ldots, X^n, \ldots)$ 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 (6) also satisfies the $CJT$, note that for $0 < \varepsilon < 1/4$,

$$Pr\left(\overline{X}_n > \frac{1}{2}\right) \geq Pr\left(\theta \geq \frac{1}{2} + 2\varepsilon\right) Pr\left(\overline{X}_n > \frac{1}{2} \mid \theta \geq \frac{1}{2} + 2\varepsilon\right). \tag{7}$$

For the second term in (7) we have:

$$\begin{aligned}
Pr\left(\overline{X}_n > \frac{1}{2} \mid \theta \geq \frac{1}{2} + 2\varepsilon\right) &= \Sigma_{k>\frac{n}{2}} Pr\left(X^1 + \cdots + X^k = k \mid \theta \geq \frac{1}{2} + 2\varepsilon\right) &(8)\\
&= \Sigma_{k>\frac{n}{2}} \binom{n}{k} \int_{\frac{1}{2}+2\varepsilon}^{1} \theta^k (1-\theta)^{n-k} dF &(9)\\
&= \int_{\frac{1}{2}+2\varepsilon}^{1} \left[\Sigma_{k>\frac{n}{2}} \binom{n}{k} \theta^k (1-\theta)^{n-k}\right] dF &(10)\\
&:= \int_{\frac{1}{2}+2\varepsilon}^{1} S_n(\theta)\, dF &(11)
\end{aligned}$$

Now, using Chebyshev's inequality we have:

$$\begin{aligned}
S_n(\theta) = Pr\left(\overline{X}_n > \frac{1}{2} \mid \theta\right) &\geq Pr\left(\overline{X}_n > \frac{1}{2} + \varepsilon \mid \theta\right) &(12)\\
&\geq 1 - \frac{V(\overline{X}_n|\theta)}{(\theta - \frac{1}{2} - \varepsilon)^2} = 1 - \frac{\theta(1-\theta)}{n(\theta - \frac{1}{2} - \varepsilon)^2} &(13)
\end{aligned}$$

Since the last expression in (13) converges to 1 uniformly on $[1/2 + 2\varepsilon, 1]$ as $n \to \infty$, taking the limit $n \to \infty$ of (11) and using (13) we have:

$$\lim_{n\to\infty} Pr\left(\overline{X}_n > \frac{1}{2} \mid \theta \geq \frac{1}{2} + 2\varepsilon\right) \geq \int_{\frac{1}{2}+2\varepsilon}^{1} dF = Pr\left(\theta \geq \frac{1}{2} + 2\varepsilon\right). \tag{14}$$

From (7) and (14) we have that for any fixed $\varepsilon > 0$,

$$\lim_{n\to\infty} Pr\left(\overline{X}_n > \frac{1}{2}\right) \geq \left[Pr\left(\theta \geq \frac{1}{2} + 2\varepsilon\right)\right]^2. \tag{15}$$

Since (15) must hold for all $1/4 > \varepsilon > 0$, and since $Pr\left(\frac{1}{2} < \theta \leq 1\right) = 1$, we conclude that

$$\lim_{n\to\infty} Pr\left(\overline{X}_n > \frac{1}{2}\right) = 1, \tag{16}$$

i.e., the sequence $X = (X^1, X^2, \ldots, X^n, \ldots)$ satisfies the $CJT$. □

To draw the consequences of Theorem 8 we prove first the following:

**Proposition 9.** *Any distribution $F$ of a variable $\theta$ in $[1/2,1]$ satisfies*

$$V(F) \leq (u - \frac{1}{2})(1-u), \tag{17}$$

*where $u = E(F)$, and equality holds in (17) only for $F$ in which*

$$Pr(\theta = \frac{1}{2}) = 2(1-u) \quad \text{and} \quad Pr(\theta = 1) = 2u - 1. \tag{18}$$

*Proof.* We want to show that

$$\int_{1/2}^{1} \theta^2 dF(\theta) - u^2 \leq (u - \frac{1}{2})(1-u), \tag{19}$$

or, equivalently,

$$\int_{1/2}^{1} \theta^2 dF(\theta) - \frac{3}{2}u + \frac{1}{2} \leq 0. \tag{20}$$

Replacing $u = \int_{1/2}^{1} \theta \, dF(\theta)$ and $\frac{1}{2} = \int_{1/2}^{1} \frac{1}{2} \, dF(\theta)$, inequality (19) is equivalent to

$$\int_{1/2}^{1} (\theta^2 - \frac{3}{2}\theta + \frac{1}{2}) \, dF(\theta) := \int_{1/2}^{1} g(\theta) \, dF(\theta) \leq 0. \tag{21}$$

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 (21). Furthermore, equality to 0 in (21) is obtained only when $F$ is such that $Pr(1/2 < \theta < 1) = 0$, and combined with $u = E(F)$ this implies (18). □

The next Proposition provides a sort of inverse to proposition 9.

**Proposition 10.** *For $(u,w) = (1,0)$ and for any pair $(u,w)$ where $1/2 < u < 1$ 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 $(u,w) = (1,0)$ the claim is trivially true (with the distribution $Pr(\theta = 1) = 1$). 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 both $\lim_{y \to u} V(F_y) = 0$ 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 $V(F_{y^*}) = w$. □

## 2.1 Presentation in the $L_2$ plane

Given a sequence $X = (X^1, X^2, ..., X^n, ...)$ of binary random variables with a joint probability distribution $P$, we define the following two parameters of $(X, P)$:

$$\underline{p} := \liminf_{n\to\infty} \overline{p}_n \tag{22}$$

$$\underline{y} := \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 \tag{23}$$

Note that this definition of $\underline{y}$ is consistent with that given in equation (5) for exchangeable variables; a case in which the limit exists.

It turns out to be useful to study the $CJT$ property of a sequence $X = (X^1, X^2, ..., X^n, ...)$ through its projection on the $(\underline{p}, \underline{y})$ plane, which we shall refer to as the $L_2$ plane. We first identify the range of this mapping:

**Proposition 11.** *For every pair $(X, P)$, the corresponding parameters $(\underline{p}, \underline{y})$ satisfy $\underline{y} \leq \underline{p}(1-\underline{p})$.*

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

$$Cov(X^i, X^j) = E(X^i X^j) - p^i p^j \leq \min(p^i, p^j) - p^i p^j.$$

Therefore,

$$E(\overline{X}_n - \overline{p}_n)^2 = \frac{1}{n^2}\left\{\sum_{i=1}^n \sum_{j\neq i} Cov(X^i, X^j) + \sum_{i=1}^n p^i(1-p^i)\right\} \tag{24}$$

$$\leq \frac{1}{n^2}\left\{\sum_{i=1}^n \sum_{j\neq i} [\min(p^i, p^j) - p^i p^j] + \sum_{i=1}^n p^i(1-p^i)\right\}. \tag{25}$$

We claim that the maximum of the last expression (25), under the condition $\sum_{i=1}^n p^i = n\overline{p}_n$, 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 (25) 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

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

□

This leads to:

**Theorem 12.** *The range of the mapping* $(X,P) \to (\underline{p},\underline{y})$ *is* (see Figure 1)

$$FE_2 = \{(u,w)|0 \le u \le 1,\ \ 0 \le w \le u(1-u)\} \tag{26}$$

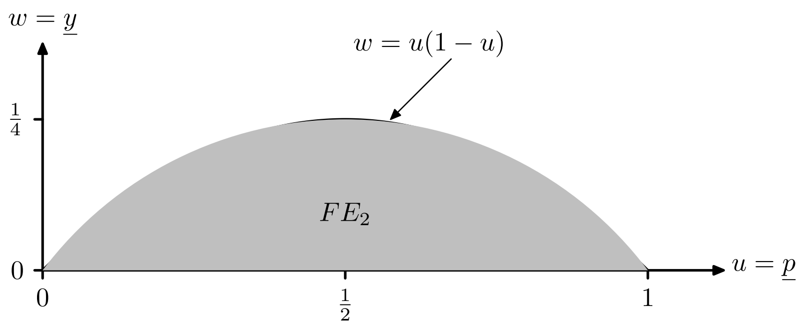


Figure 1: The feasible set $FE_2$

That is, for any pair $(X,P)$, we have $(\underline{p},\underline{y}) \in FE_2$ and for any $(u,w) \in FE_2$ there is a pair $(X,P)$ for which $\underline{p} = u$ and $\underline{y} = w$.

*Proof.* The first part follows from Proposition 11 (since clearly $\underline{y} \ge 0$). For the second part, observe first, as we have remarked in the proof of Proposition 11, that for the pair $(X,P)$ in which $P\{X^1 = X^2 = \ldots = 1\} = u$ and $P\{X^1 = X^2 = \ldots = 0\} = 1-u$ we have $p^1 = p^2 = \cdots = p^n = \overline{p}_n = u$ and hence $\underline{p} = u$. Also, for all $n = 1,2,\ldots$,

$$E(\overline{X}_n - \overline{p}_n)^2 = E(\overline{X}_n - u)^2 = u(1-u) \quad \text{and hence} \quad \underline{y} = \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 = u(1-u),$$

which means that any point on the parabola $w = u(1-u)$ is attainable as an image of a pair $(X,P)$. Next note that for $u \in [0,1]$, the pair $(Y,\tilde{P})$ in which $(Y_i)_{i=1}^{\infty}$ are i.i.d. with $\tilde{P}\{Y_i = 1\} = u$ and $\tilde{P}\{Y_i = 0\} = 1-u$ is mapped to $(\underline{p},\underline{y}) = (u,0)$ since

$$\underline{y} = \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2 = \liminf_{n\to\infty} \frac{1}{n^2}\Sigma_{i=1}^{n} E(\overline{X}_i - u)^2 = \liminf_{n\to\infty} \frac{1}{n^2}\Sigma_{i=1}^{n} u(1-u) = \liminf_{n\to\infty} \frac{u(1-u)}{n} = 0.$$

It remains to prove that all interior points of $FE_2$ are attainable. Let $(u,w)$ be such an interior point, that is, $0 < u < 1$ and $0 < w < u(1-u)$. Define the pair $(Z,Q)$ to be the above-defined pair $(X,P)$ with probability $w/u(1-u)$ and the above-defined $(Y,\tilde{P})$ with probability $1 - w/u(1-u)$. It is readily seen that this pair is mapped to

$$\frac{w}{u(1-u)}(u,u(1-u)) + \left(1 - \frac{w}{u(1-u)}\right)(u,0) = (u,w).$$

□

The geometric expression of Theorem 8, combined with Theorem 6, Proposition 9, and Proposition 10, can now be stated as follows: In the $L_2$ plane of $(\underline{p},\underline{y})$ let

$$A = \left\{ (\underline{p},\underline{y}) \;\middle|\; \frac{1}{2} < \underline{p} \leq 1; \text{ and } \underline{y} < (\underline{p} - \frac{1}{2})(1-\underline{p}) \right\} \bigcup \{(1,0)\} \tag{27}$$

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

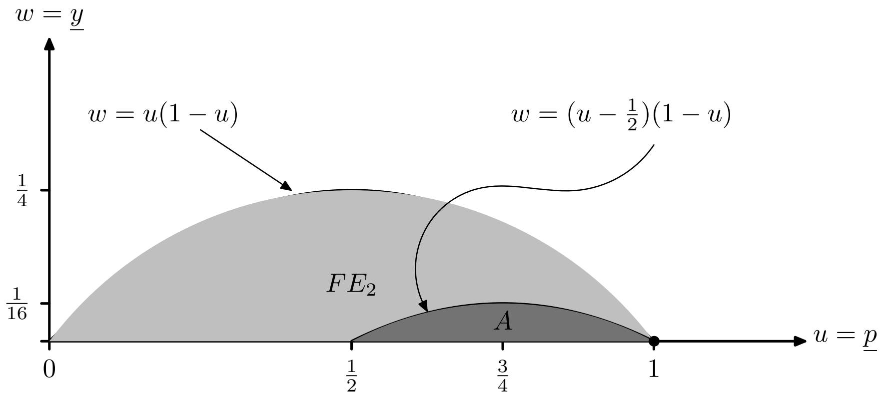


Figure 2: The $CJT$ region for exchangeable variables.

**Theorem 13.** 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 6, Theorem 8, and Proposition 9.
- Part 2. follows de Finetti's Theorem 6, Theorem 8, and Proposition 10.

□

## 2.2 Application to symmetric juries

A jury game $G_n$, as defined in Section 1, is said to be *symmetric* if

- $T^1 = T^2 = \ldots = T^n$
- The probability distribution $p^{(n)}$ is symmetric in the variables $t^1, \ldots, t^n$.

We consider a sequence of increasing juries $(G_n)_{n=1}^{\infty}$ such that $G_n$ is symmetric for all $n$. In such a sequence $\Sigma_n^i$ is the same for all $i$ and all $n$ and is denoted by $\Sigma$. A strategy vector $\sigma_n = (\sigma_n^1, \dots, \sigma_n^n) \in \Sigma_n$ is said to be *symmetric*, if $\sigma_n^1 = \sigma_n^2 = \dots = \sigma_n^n$.

**Corollary 14.** *Let* $\underset{\sim}{\sigma} = (\sigma, \sigma, \dots, \sigma, \dots) \in \Sigma^\infty$ *and let* $\underset{\sim}{X} = (X^1, X^2, \dots, X^n, \dots)$ *be the sequence of binary random variables derived from* $\underset{\sim}{\sigma}$ *by* (1)*; then* $\underset{\sim}{X}$ *is exchangeable. If* $\underset{\sim}{X}$ *satisfies* (6)*, then there exists a sequence of BNE,* $\sigma_*^n = (\sigma_n^*, \dots, \sigma_n^*)$ *of* $G_n$ *for* $n = 1, 2, \dots$*, that satisfies the CJT.*

*Proof.* Follows from Theorem 8 and Theorem 2 of McLennan (1998). □

# 3 Sufficient conditions

Having characterized the *CJT* conditions for exchangeable variables we proceed now to the general case and we start with sufficient conditions.

Let $X = (X^1, X^2, \dots, X^n, \dots)$ 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

$$\lim_{n\to\infty} P\left(\Sigma_{i=1}^n X^i > \frac{n}{2}\right) = 1 \tag{28}$$

This is the condition corresponding to condition (2) (on page 5) when $X_i = X_i(\sigma_i)$ for an infinite sequence of constant strategies $(\sigma^i)_{i=1}^\infty$ that satisfy *CJT*.

In this section we provide sufficient conditions for a pair $(X, P)$ to satisfy the *CJT*. Recall our notation: $\overline{X}_n = (X^1 + X^2, \dots + X^n)/n$, $p^i = E(X^i)$ and $\overline{p}_n = (p^1 + p^2, \dots + p^n)/n$.

**Theorem 15.** *Assume that* $\overline{p}_n > \frac{1}{2}$ *for all* $n > N_0$ *and*

$$\lim_{n\to\infty} \frac{E(\overline{X}_n - \overline{p}_n)^2}{(\overline{p}_n - \frac{1}{2})^2} = 0, \tag{29}$$

*or equivalently assume that*

$$\lim_{n\to\infty} \frac{\overline{p}_n - \frac{1}{2}}{\sqrt{E(\overline{X}_n - \overline{p}_n)^2}} = \infty; \tag{30}$$

*then the CJT is satisfied.*

*Proof.*

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

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

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

As this last term tends to zero by (29), the $CJT$ (28) then follows. □

**Corollary 16.** *If* $\Sigma_{i=1}^{n} \Sigma_{j \neq i} Cov(X^i, X^j) \leq 0$ *for* $n > N_0$ *(in particular if* $Cov(X^i, X^j) \leq 0$ *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
$p(1-p) \leq 1/4$ we have for $n > N_0$,

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

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

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

□

**Remark 17.** *It follows from equation (29) that any* $(X, P)$ *satisfying this sufficient condition must have* $\underline{y} = 0$*; that is, it corresponds to a point* $(\underline{p}, 0)$ *in the* $L_2$ *plane. Thus, any distribution with* $\underline{y} > 0$ *that satisfy the* $CJT$*, does not satisfy this sufficient condition. In particular, this is true for the exchangeable sequences (with* $\underline{y} > 0$*) we identified in Section 2 and the non-exchangeable sequences satisfying the* $CJT$ *we will see in Section 6.*

**Remark 18.** *Note that under the condition of corollary 16, namely, for bounded random variables with all covariances being non-positive, the (weak) Law of Large Numbers (LLN) holds for arbitrarily dependent variables (see, e.g., Feller (1957), Vol. I, exercise 9, p. 262). This is not implied by corollary 16 since, as we show in Appendix 8.3, the* $CJT$*, strictly speaking, is not a Law of Large Lumbers. In particular,* $CJT$ *does not imply LLN and LLN does not imply* $CJT$*.*

**Remark 19.** *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)).*

# 4 Necessary conditions

We start this section with a simple observation and then state two necessary conditions that do not fully imply one another in either direction.

**Proposition 20.** *Given a sequence* $X = (X^1, X^2, \ldots, X^n, \ldots)$ *of binary random variables with a joint probability distribution* $P$*, if the* $CJT$ *holds then* $\underline{p} \geq \frac{1}{2}$*.*

*Proof.* Define a sequence of events $(B_n)_{n=1}^{\infty}$ by $B_n = \{\omega \mid \overline{X}_n(\omega) \geq 1/2\}$. 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

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

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

## 4.1 A necessary condition in the $L_2$ plane

In this subsection we provide a necessary condition in $L_2$ for a general sequence $(X,P)$ to satisfy the $CJT$. That is, a condition in terms of two characteristics, $\underline{p} = \liminf_{n\to\infty} \overline{p}_n$ and $\underline{y} = \liminf_{n\to\infty} E(\overline{X}_n - \overline{p}_n)^2$.

**Theorem 21.** *Let* $X = (X^1, X^2, \ldots, X^n, \ldots)$ *be a sequence of binary random variables with joint distribution* $P$*. If* $(X,P)$ *satisfy the* $CJT$*, then* $\underline{y} \leq (\underline{p} - \frac{1}{2})(1 - \underline{p})$*.*

*Proof.* Recall our notation $B_n = \{\omega \in \Omega \mid \overline{X}_n(\omega) \geq \frac{1}{2}\}$; then, since $(X,P)$ satisfy the $CJT$, $\lim_{n\to\infty} P(B_n) = 1$. The main part of the proof is a direct computation of $E(\overline{X}_n(\omega) - \overline{p}_n)^2$. Denote by $B_n^c := \Omega \setminus B_n$ the complement of $B_n$; then:

$$\begin{aligned}
E(\overline{X}_n(\omega) - \overline{p}_n)^2 &= E\left(\overline{X}_n(\omega) - \frac{1}{2} + \frac{1}{2} - \overline{p}_n\right)^2 \\
&= E\left(\overline{X}_n(\omega) - \frac{1}{2}\right)^2 + 2\left(\frac{1}{2} - \overline{p}_n\right) E\left(\overline{X}_n(\omega) - \frac{1}{2}\right) + \left(\frac{1}{2} - \overline{p}_n\right)^2 \\
&= E\left(\overline{X}_n(\omega) - \frac{1}{2}\right)^2 - \left(\frac{1}{2} - \overline{p}_n\right)^2 \\
&= \int_{B_n^c} \left(\overline{X}_n(\omega) - \frac{1}{2}\right)^2 dP + \int_{B_n} \left(\overline{X}_n(\omega) - \frac{1}{2}\right)\left(\overline{X}_n(\omega) - \frac{1}{2}\right) dP - \left(\frac{1}{2} - \overline{p}_n\right)^2 \\
&\leq \int_{B_n^c} \left(\overline{X}_n(\omega) - \frac{1}{2}\right)^2 dP + \frac{1}{2}\int_{B_n} \left(\overline{X}_n(\omega) - \frac{1}{2}\right) dP - \left(\frac{1}{2} - \overline{p}_n\right)^2.
\end{aligned}$$

Thus,

$$\begin{aligned} E(\overline{X}_n(\omega)-\overline{p}_n)^2 &\leq \int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)^2 dP - \frac{1}{2}\int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)dP + \frac{1}{2}E\left(\overline{X}_n(\omega)-\frac{1}{2}\right) - \left(\frac{1}{2}-\overline{p}_n\right)^2 \\ &= \frac{1}{2}\left(\overline{p}_n-\frac{1}{2}\right) - \left(\frac{1}{2}-\overline{p}_n\right)^2 + \int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)^2 dP - \frac{1}{2}\int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)dP \\ &= \left(\overline{p}_n-\frac{1}{2}\right)(1-\overline{p}_n) + \int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)^2 dP - \frac{1}{2}\int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)dP \end{aligned}$$

For any $\varepsilon > 0$ there exists $N(\varepsilon)$ such that for $n > N(\varepsilon)$,

$$\int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)^2 dP < \frac{\varepsilon}{2} \quad \text{and} \quad \left|\frac{1}{2}\int_{B_n^c}\left(\overline{X}_n(\omega)-\frac{1}{2}\right)dP\right| < \frac{\varepsilon}{2}.$$

Hence for $n > N(\varepsilon)$,

$$E(\overline{X}_n(\omega)-\overline{p}_n)^2 \leq \left(\overline{p}_n-\frac{1}{2}\right)(1-\overline{p}_n)+\varepsilon. \tag{31}$$

We conclude that

$$\underline{y} = \liminf_{n\to\infty} E(\overline{X}_n-\overline{p}_n)^2 \leq \liminf_{n\to\infty}\left(\overline{p}_n-\frac{1}{2}\right)(1-\overline{p}_n)+\varepsilon,$$

for every $\varepsilon > 0$. Hence

$$\underline{y} = \liminf_{n\to\infty} E(\overline{X}_n-\overline{p}_n)^2 \leq \liminf_{n\to\infty}\left(\overline{p}_n-\frac{1}{2}\right)(1-\overline{p}_n).$$

Choose a sequence $(n_k)_{k=1}^{\infty}$ such that $\lim_{k\to\infty}\overline{p}_{n_k} = \underline{p}$; then

$$\underline{y} \leq \liminf_{k\to\infty}\left(\overline{p}_{n_k}-\frac{1}{2}\right)(1-\overline{p}_{n_k}) = \left(\underline{p}-\frac{1}{2}\right)(1-\underline{p}).$$ [6]

□

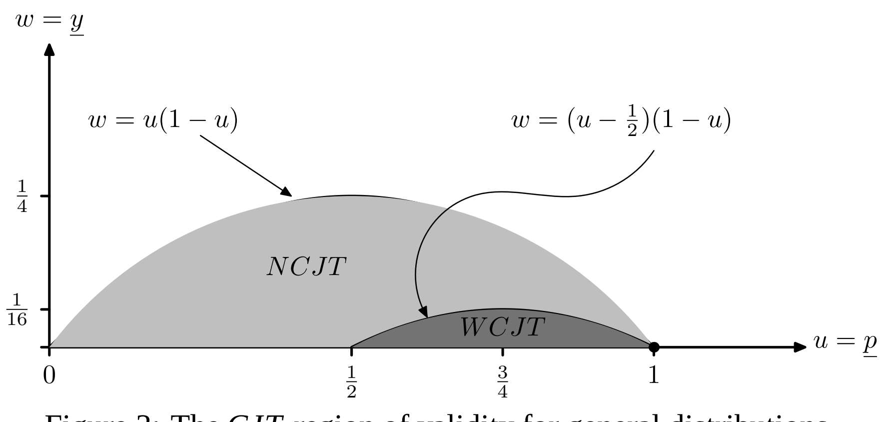


Figure 3: The $CJT$ region of validity for general distributions.

[6]Since for any $\varepsilon > 0$ inequality (31) holds for all $n > N(\varepsilon)$, then for a subsequence $(n_k)_{k=1}^{\infty}$ such that $\lim_{k\to\infty}\overline{p}_{n_k} = \tilde{p}$ and $\lim_{k\to\infty}E(\overline{X}_{n_k}-\overline{p}_{n_k})^2 = \tilde{y}$, we get $\tilde{y} \leq (\tilde{p}-\frac{1}{2})(1-\tilde{p})$. It follows that if $(X,P)$ satisfies the $CJT$, then any limit point of $(\overline{p}_n, E(\overline{X}_n-\overline{p}_n)^2)$ is in the region $A$ of Figure 2 (or region $WCJT$ in Figure 3). We are indebted to A. Neyman for a discussion concerning this observation.

Figure 3 depicts the regions of validity of the $CJT$ in the $L_2$ plane: Any distribution for which the parameters $(\underline{p},\underline{y})$ lie in the lightly colored region denoted by $NCJT$ (for *not* $CJT$) does not satisfy the $CJT$. The dark region, denoted by $WCJT$ (for *weak* $CJT$), is the closed area below the small parabola. Any distribution that satisfies the $CJT$ must have parameters $(\underline{p},\underline{y})$ in this region. As we saw in Section 2, for exchangeable random variables, the region $WCJT$ (excluding the parabola and including the point $(1,0)$) defines also a sufficient condition: Any sequence of exchangeable variables whose parameters $(\underline{p},\underline{y})$ lie in this region satisfy the $CJT$. However, for general distributions this is not a sufficient condition; as we shall see later, for any $(\underline{p},\underline{y})$ in this region, excluding $(1,0)$, there is a sequence with these parameters that does not satisfy the $CJT$.

## 4.2 A necessary condition in the $L_1$ plane

In this subsection we provide a necessary condition in $L_1$ for a general sequence $(X,P)$ to satisfy the $CJT$. That is, a condition in terms of two characteristics, $\underline{p} = \liminf_{n\to\infty} \overline{p}_n$ and $\underline{y}^* = \liminf_{n\to\infty} E\left|\overline{X}_n - \overline{p}_n\right|$.

**Theorem 22.** *Let $X = (X^1, X^2, \ldots, X^n, \ldots)$ be a sequence of binary random variables with joint distribution P. If $(X,P)$ satisfy the CJT, then $\underline{y}^* \leq 2(2\underline{p}-1)(1-\underline{p})$.*

*Proof.* See Appendix 7.2

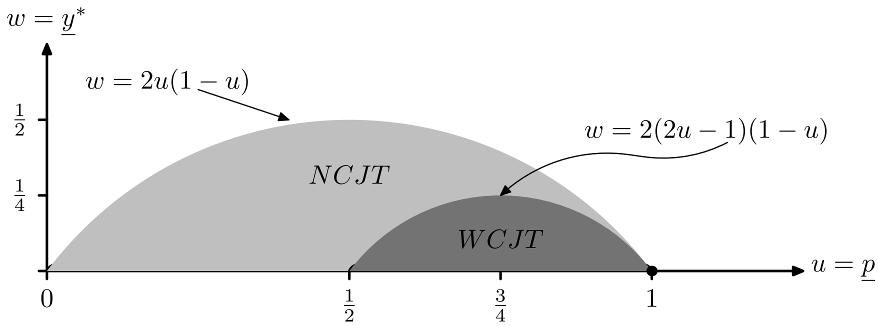


Figure 4: The $CJT$ region of validity in $L_1$.

Figure 4 depicts the regions of validity of the $CJT$ in the $L_1$ plane; the analogue of Figure 3.

Strangely enough, Theorem 22 and Theorem 21 do not imply each other in either direction. Furthermore, the techniques of the proofs $L_1$ and in $L_2$ are very different. We could derive only a weak implication in one direction which stems from the following lemma:

**Lemma 23.** *One always has: $\underline{y}^* \geq 2\underline{y}$.*

*Proof.* Denoting $A_n = \{\omega \in \Omega \mid \overline{p}_n - \overline{X}_n(\omega) \geq 0\}$, we have:

$$\begin{aligned}\int_{A_n} (\overline{p}_n - \overline{X}_n)^2 \, dP &= \int_{A_n} (\overline{p}_n - \overline{X}_n)(\overline{p}_n - \overline{X}_n) \, dP \\ &\leq \overline{p}_n \int_{A_n} (\overline{p}_n - \overline{X}_n) \, dP = \overline{p}_n \frac{y_n^*}{2}.\end{aligned}$$

Similarly,

$$\int_{A_n^c} (\overline{X}_n - \overline{p}_n)^2 \, dP \leq (1 - \overline{p}_n) \frac{y_n^*}{2}.$$

Hence for all $n$ we have:

$$y_n := E(\overline{X}_n - \overline{p}_n)^2 = \int_{\Omega} (\overline{X}_n - \overline{p}_n)^2 \, dP \leq \overline{p}_n \frac{y_n^*}{2} + (1 - \overline{p}_n) \frac{y_n^*}{2} = \frac{y_n^*}{2}.$$

Taking a subsequence $(n_k)_{k=1}^{\infty}$ such that $\lim_{k\to\infty} y_{n_k}^* = \underline{y}^*$, we conclude that

$$\underline{y}^* \geq 2 \liminf_{k\to\infty} y_{n_k} \geq 2\underline{y}.$$

□

Combining Lemma 23 with Theorem 22 yields,

**Corollary 24.** *Let $X = (X^1, X^2, \ldots, X^n, \ldots)$ be a sequence of binary random variables with joint distribution $P$. If $\underline{y} > (2\underline{p} - 1)(1 - \underline{p})$, then $(X, P)$ does not satisfy the CJT.*

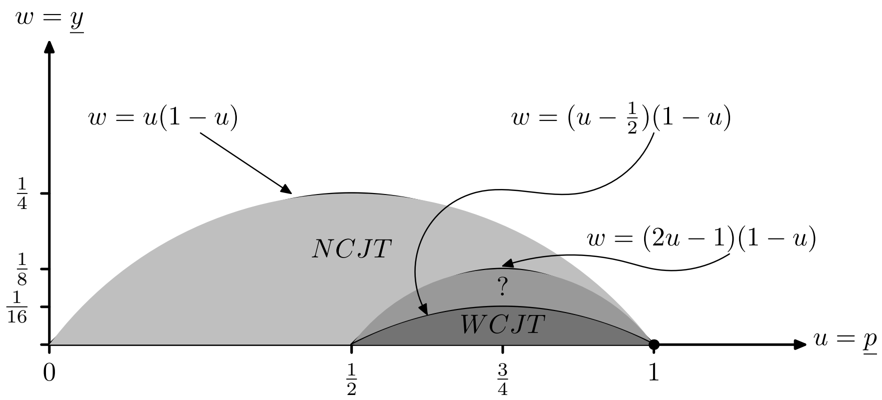


Figure 5: The $CJT$ validity region in $L_2$ as implied by the condition in $L_1$.

Figure 5 depicts the conclusion of the last corollary: The region with the lightest color, denoted by $NCJT$, is the region in which the $CJT$ is not satisfied for any $(X, P)$ with these values of $(\underline{p}, \underline{y})$. The darkest region, denoted by $WCJT$, is the region of $(\underline{p}, \underline{y})$ for which there exist $(X, P)$ with these parameters that satisfy the $CJT$. Clearly, this is a weaker result than Theorem 21 that we obtained directly in $L_2$ and is described in Figure 3 according to which, the crescent in Figure 5, denoted by "?", belongs to the $NCJT$ region.

# 5 Distributions in $WCJT$ that do not satisfy the $CJT$

In this section we prove that the necessary conditions stated in Theorems 21 and 22 are not sufficient. In fact we prove a stronger result, namely: To any pair of parameters in the closure of the dark $WCJT$ region (either in Figure 3 in $L_2$ or in Figure 4 in $L_1$), excluding the point $(1, 0)$, there is a distribution that does not satisfy the $CJT$. We shall prove this only for the $L_2$ plane (the proof for the $L_1$ plane is similar). This is established by the following:

**Theorem 25.** *For any* $(u,w) \in \{(u,w) \mid 0 < u < 1\,;\;\; 0 \leq w \leq u(1-u)\}$, *there is a sequence of binary random variables* $Z$ *with joint distribution* $H$ *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 $0 < t \leq 1$ let $(Z_t,H_t)$ be the pair in which $Z_t^i = tX^i + (1-t)Y^i$ 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 $E(Z_t^i) = u$ 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 $0 < t \leq 1$.

Finally, $(Z_t,H_t)$ does not satisfy the $CJT$ since for all $n$,

$$Pr(\overline{Z}_{t,n} > \frac{1}{2}) \leq 1 - Pr(Z_t^1 = Z_t^2 = \ldots = 0) = 1 - t(1-u) < 1.$$

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 $B_\ell$ the set of indices of the $\ell$-th block and its cardinality by $b_\ell$. Thus $n(\ell) = \Sigma_{j=1}^{\ell} b_j$ is the index of $Z^i$ at the end of the $\ell$-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 $b_\ell$ 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,

$$Z^i = \begin{cases} X^i & \text{if } \ i \in B_{2k-1} & \text{for some } \ k = 1, 2, \ldots \\ Y^i & \text{if } \ i \in B_{2k} & \text{for some } \ k = 1, 2, \ldots \end{cases}$$

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 $\ell$-th block and by $n(\ell) = n_x(\ell) + n_y(\ell)$ the number of coordinates at the end of the $\ell$-th block of $Z$. It follows from 1 and 2 (in the definition of $b_\ell$) 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 theorem. □

An immediate implication of Theorem 25 is the following:

**Corollary 26.** *For any pair of parameters* $(\underline{p}, \underline{y})$ *satisfying* $1/2 \leq \underline{p} < 1$ *and* $0 \leq \underline{y} \leq (\underline{p}-1/2)(1-\underline{p})$ (that is, the point $(\underline{p}, \underline{y})$ is in the closure of the region $WCJT$ in Figure 3, excluding $(1,0)$)*, there is a distribution with these parameters that does not satisfy the* $CJT$.

# 6 Non-exchangeable sequences satisfying the $CJT$

In this section prove the existence of sequences $(X,P)$ of dependent random variables, sequences that are non-exchangeable and satisfy the $CJT$. By Theorem 21, such distributions must have their parameter in the closure of the dark $WCJT$ region (either in Figure 3 in $L_2$ or in Figure 4 in $L_1$). In fact, we shall prove that for any point in this region there is a distribution that satisfies the $CJT$, and is not exchangeable. We shall prove that only in the $L_2$ plane. The proof for the $L_1$ plane is similar. The construction of these sequences uses the idea of the *interlacing* of two sequences, which can be generalized and proves to be useful.

**Theorem 27.** *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* $CJT$.

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

$$Z = (Y^1, Y^2, X^1, Y^3, X^2, Y^4, X^3, ..., Y^n, X^{n-1}, Y^{n+1}, X^n ...),$$

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 $\underline{\tilde{y}} = \frac{1}{4}\underline{y}$. 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 $\underset{\sim}{\tilde{p}} = 1 - t + t\underline{p}$ and $\underset{\sim}{\tilde{y}} = t^2\underline{y}$ is verified by straightforward computation. □

**Remark 28.**

- *Note that the sequence $Z$ is clearly not exchangeable* (except for the case $t = 0$ which corresponds to) $(1,0)$).
- *The interlacing of the two sequences $X$ and $Y$ described in the proof of Theorem 27 may be defined for any $t \in [0,1]$. We were specifically interested in $t \in [0,1/2]$ since this guarantees the $CJT$.*

Figure 6 depicts the interlacing procedure: The parabolic line joining $(u^*, w^*)$ to the point $(1,0)$, corresponds to all interlacing with $t \in [0,1]$. The lower part, described as a thick line, corresponds to interlacing when $t \in [0,1/2]$. For these values of $t$, the interlacing yields distributions satisfying the $CJT$. The small parabola is the locus of points corresponding to $t = 1/2$ when $(u^*, w^*)$ ranges over the parabola $w = u(1-u)$.

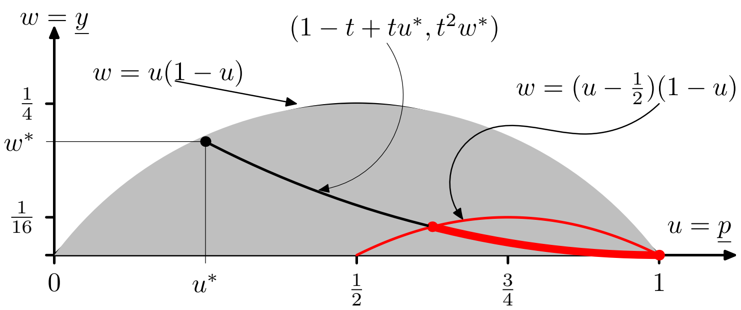


Figure 6: Interlacing with $(1,0)$ in $L_2$.

**Corollary 29.** *For any $(\underline{p}, \underline{y})$ in the set*

$$\overline{A} = \{(\underline{p}, \underline{y}) \mid 0 \leq \underline{y} \leq (\underline{p} - 1/2)(1 - \underline{p}) \ ; \quad 1/2 \leq \underline{p} \leq 1\}$$

(this is the closure of the region $WCJT$ in Figure 3), *there is a sequence of non-exchangeable random variables, with these parameters, that satisfy the $CJT$.*

*Proof.* By straightforward verification we observe that the set $\overline{A}$ is obtained from Theorem 27 by letting $(\underline{p},\underline{y})$ range over the points of parabola $w = u(1-u)$ defining the feasible set $FE_2$. In other words, $\overline{A}$ can also be written as:

$$\overline{A} = \{(\underline{p},\underline{y}) \mid \underline{p} = 1-t+tu\ ;\ \underline{y} = t^2u(1-u)\ ;\ 0 \le t \le 1/2,\ 0 \le u \le 1\}$$

□

Note that $\overline{A}$ is the closure of the set $A$ defined in equation (27) for exchangeable variables, but $\overline{A} \neq A$. More specifically, the points $(\underline{p},\underline{y})$ on the parabola $\underline{y} = (\underline{p}-1/2)(1-\underline{p})$, excluding $(1,0)$, are in $\overline{A}$ but not in $A$. For each of these points there is a corresponding sequence satisfying the $CJT$ but this sequence cannot be exchangeable.

Finally, combining Corollary 29 and Theorem 25 yields:

**Corollary 30.** *For any point $(\underline{p},\underline{y})$ in $\overline{A}\setminus\{(1,0)\}$ there is a corresponding sequence satisfying the $CJT$ and a corresponding sequences that does not satisfy the $CJT$.*

## 6.1 Other distributions satisfying the $CJT$: General interlacing

So far we have identified three types of distributions that satisfy the $CJT$; all correspond to parameters $(\underline{p},\underline{y})$ in the set $\overline{A}$, the closure of the region $WCJT$ in Figure 3.

1. Distributions satisfying the sufficient condition (Theorem 15).

2. Exchangeable distributions characterized in Theorem 8.

3. Non-exchangeable distributions obtained by interlacing with constant sequence $Y = (1,1,\ldots)$ (Theorem 27).

In this section we construct more distributions satisfying the $CJT$ that are not in either of the three families mentioned above. We do so by generalizing the notion of the "interlacing" of two distributions that we introduced in Section 6.

**Definition 31.** *Let $X = (X^1, X^2, \ldots, X^n, \ldots)$ be a sequence of binary random variables with joint probability distribution $F$ and let $Y = (Y^1, Y^2, \ldots, Y^n, \ldots)$ 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 $n = 1,2,\ldots$,*

$$Z^n = \begin{cases} X^{\lfloor tn \rfloor} & \text{if} \quad \lfloor tn \rfloor > \lfloor t(n-1) \rfloor \\ Y^{\lceil (1-t)n \rceil} & \text{if} \quad \lceil (1-t)n \rceil > \lceil (1-t)(n-1) \rceil \end{cases}, \tag{34}$$

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

The following lemma is a direct consequence of Definition 31.

**Lemma 32.** *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. □

Thus, from any two distributions satisfying the $CJT$ we can construct a continuum of distributions satisfying the $CJT$. These distributions will generally be outside the union of the three families listed above.

# 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.

Ben-Yashar, R. and J. Paroush (2000), "A non-asymptotic Condorcet jury theorem". Social Choice and Welfare 17: 189–199.

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

Berend, D. and L. Sapir (2007), "Monotonicity in Condorcet's jury theorem with dependent voters". Social Choice and Welfare 28: 507–528.

Berg, S. (1993a), "Condorcet's jury theorem, dependency among jurors". Social Choice and Welfare 10: 87–96.

Berg, S. (1993b), "Condorcet's jury theorem revisited". European Journal of Political Economy 9: 437–446.

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

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.

Dietrich, F. (2008), "The premises of Condorcet's jury theorem are not simultaneously justified". To appear in Episteme.

Dietrich, F. and C. List (2004), "A model of jury decision where all the jurors have the same evidence". Synthese 142: 175–202.

Duggan, J. and C. Martinelli (2001), “A Bayesian model of voting in juries”. Games and Economic Behavior 37: 259–294.

Estlund, D. (1994), “Opinion leaders, independence and Condorcet’s jury theorem”. Theory and Decision 36: 131–162.

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.

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

Ladha, K. K. (1992), “The Condorcet jury theorem, free speech and correlated votes”. American Journal of Political Science 36: 617–634.

Ladha, K. K. (1993), “Condorcet’s jury theorem in the light of de Finetti’s theorem: Majority-voting with correlated votes”. Social Choice and Welfare 10: 69–85.

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.

McLennan, A. (1998), “Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents”. American Political Science Review 92: 413–419.

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 (1982), “Optimal decision rules in dichotomous choice situations”. International Economic Review 23: 289–297.

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.

# 7 Appendix

## 7.1 Every sequence of of binary random variables is attainable

In this section we prove what we claimed on page 7, namely, that for any infinite sequence of binary random variables $X$ there is a sequence of games $(G_n)_{n=1}^{\infty}$ and an infinite sequence of constant strategies $\sigma = (\sigma^1, \sigma^2, \ldots, \sigma^n, \ldots)$ that yield this $X$ as the infinite sequence of the indicators of correct voting.

Let $X = (X^1, X^2, ..., X^n, ...)$ be a sequence of binary random variables on some probability space $(\tilde{\Omega}, \mathscr{B}, \mathscr{P})$. Let $P$ also denote the distribution of $X$. In our model let $T^i = \{t_0^i, t_1^i\}$ be the type set of juror $i$ and let the type of juror $i$, $t^i = t^i(\theta, X^i(\omega))$ be defined by: $t^i(g,0) = t^i(z,1) = t_0^i$ and $t^i(g,1) = t^i(z,0) = t_1^i$. We define the probability distribution $p^{(n)}$ on $\Omega_n = \Theta \times T^1 \times \ldots \times T^n$ as follows: Let $p^{(n)}(z) = p^{(n)}(g) = 1/2$; for $\varepsilon_k \in \{0,1\}$; $\quad k = 1, \ldots, n$ let

$$\tilde{p}(g, X^1 = \varepsilon_1, \ldots, X^n = \varepsilon_n) = \tilde{p}(z, X^1 = 1 - \varepsilon_1, \ldots, X^n = 1 - \varepsilon_n) = \frac{1}{2} P(X^1 = \varepsilon_1, \ldots, X^n = \varepsilon_n)$$

and define

$$p^{(n)}(\theta, t^1, t^2, \ldots, t^n) = \tilde{p}(\theta, X^1, X^2, \ldots, X^n).$$

The sequence $(p^{(n)})_{n=1}^{\infty}$ clearly satisfies the projective structure required for the Kolmogorov's extension theorem (that is, the marginal distribution of $p^{(n+1)}$ on $\Omega_n$ is equal to $p^{(n)}$). It defines therefore a probability distribution $p$ on $\Omega = \lim_{\infty \leftarrow n} \Omega_n$.

Define now the (informative voting) strategies $\sigma^i$ by: $\sigma^i(t_0^i) = a$ and $\sigma^i(t_1^i) = c$, and let $\tilde{X}^1, \ldots, \tilde{X}^n \ldots$ be the indicators of correct voting (w.r.t. this $\sigma$), then

$$\tilde{X}^i(g, t^i(g,1)) = \tilde{X}^i(z, t^i(z,1)) = 1 \quad \text{and} \quad \tilde{X}^i = 0 \;\; \text{otherwise}.$$

Thus

$$\tilde{X}^i(\theta, \omega) = 1 \quad \Leftrightarrow \quad X^i(\omega) = 1,$$

which means that we obtained the original given sequence.

## 7.2 Proof of Theorem 22

In this section we provide a necessary condition for a general sequence of binary random variables $X = (X^1, X^2, ..., X^n, ...)$ with joint distribution $P$, in terms of two of its characteristics namely, $\underline{p} = \liminf_{n\to\infty} \overline{p}_n$ and $\underline{y}^* = \liminf_{n\to\infty} E|\overline{X}_n - \overline{p}_n|$.

Let $y_n^* = E|\overline{X}_n - \overline{p}_n|$; then $\underline{y}^* = \liminf_{n\to\infty} y_n^*$. For $n = 1, 2, \ldots$, let

$$A_n = \{\omega \in \Omega \mid \overline{p}_n - \overline{X}_n(\omega) \geq 0\} \quad \text{and} \quad A_n^c = \Omega \setminus A_n.$$

Then, since $E(\overline{X}_n - \overline{p}_n) = 0$,

$$\int_{A_n^c} (\overline{X}_n - \overline{p}_n)\, dP = \frac{y_n^*}{2} \;\text{ and }\; \int_{A_n^c} (1 - \overline{p}_n)\, dP = (1 - \overline{p}_n) P(A_n^c) \geq \frac{y_n^*}{2}. \text{ Hence,}$$

$$P(A_n) = 1 - P(A_n^c) \leq 1 - \frac{y_n^*}{2(1-\overline{p}_n).} \tag{35}$$

Also,

$$\int_{A_n} (\overline{p}_n - \overline{X}_n)\, dP = \overline{p}_n P(A_n) - \int_{A_n} \overline{X}_n\, dP = \frac{y_n^*}{2}. \tag{36}$$

Hence, since $\overline{X}_n \geq 0$,

$$P(A_n) \geq \frac{y_n^*}{2\overline{p}_n.} \tag{37}$$

Assuming $y^* > 0$ and $\underline{p} < 1$, it follows from (37) and (35) that there is a subsequence $(n_k)_{k=1}^{\infty}$ such that $(P(A_{n_k}))_{k=1}^{\infty}$ is uniformly bounded away from 0 and 1,

$$\lim_{k\to\infty} \overline{p}_{n_k} = \underline{p} \quad \text{and} \quad \lim_{k\to\infty} P(A_{n_k}) = \ell \quad \text{where} \quad 0 < \ell < 1. \tag{38}$$

**Lemma 33.** *Let $t > 0$; then*

$$\liminf_{k\to\infty} P\left(\left\{\omega \in A_{n_k} \,\middle|\, \overline{p}_{n_k} - \overline{X}_{n_k}(\omega) \leq \frac{y_{n_k}^*}{2P(A_{n_k})} - t\right\}\right) < \ell. \tag{39}$$

*Proof.* Assume by contradiction that (39) does not hold; then, since the sets on the left-hand side are subsets of $A_{n_k}$, it follows from (38) that:

$$\lim_{k\to\infty} P\left(\left\{\omega \in A_{n_k} \,\middle|\, \overline{p}_{n_k} - \overline{X}_{n_k}(\omega) \ \leq \ \frac{y_{n_k}^*}{2P(A_{n_k})} - t\right\}\right) = \ell. \tag{40}$$

Denote: $\tilde{A}_{n_k} = \left\{\omega \in A_{n_k} \,\middle|\, \overline{p}_{n_k} - \overline{X}_{n_k}(\omega) \ \leq \ \frac{y_{n_k}^*}{2P(A_{n_k})} - t\right\}$. Then,

$$\overline{p}_{n_k} P(\tilde{A}_{n_k}) - \int_{\tilde{A}_{n_k}} \overline{X}_{n_k}(\omega) dP \ \leq \ \frac{y_{n_k}^* P(\tilde{A}_{n_k})}{2P(A_{n_k})} - tP(\tilde{A}_{n_k}). \tag{41}$$

Clearly, $\lim_{k\to\infty} P(\tilde{A}_{n_k}) = \ell = \lim_{k\to\infty} P(A_{n_k})$ and since $\tilde{A}_{n_k} \subseteq A_{n_k}$, we have

$$\lim_{k\to\infty} \left| \int_{\tilde{A}_{n_k}} \overline{X}_{n_k}(\omega) dP - \int_{A_{n_k}} \overline{X}_{n_k}(\omega) dP \right| = 0.$$

Thus for $k_0$ sufficiently large, the inequality (41) contradicts the last equality in (36) for $n = n_{k_0}$. $\square$

Let

$$B_n = A_n \setminus \tilde{A}_n = \left\{\omega \in A_n \,\middle|\, \overline{p}_n - \overline{X}_n(\omega) \ > \ \frac{y_n^*}{2P(A_n)} - t\right\};$$

then, by Lemma 33, there is a subsequence $(B_{n_k})_{k=1}^{\infty}$ and $q > 0$, such that $P(B_{n_k}) > q > 0$ for all $k$, that is

$$\overline{X}_{n_k}(\omega) < \overline{p}_{n_k} - \frac{y_{n_k}^*}{2P(A_{n_k})} + t; \quad \forall \omega \in B_{n_k}; \ \forall k. \tag{42}$$

**Example 34.** *Let* $\frac{1}{2} \leq \underline{p} < 1$ *and* $\underline{y}^* = 2\underline{p}(1-\underline{p})$*; then, by (35) and (42) we have*

$$\overline{X}_{n_k}(\omega) < \overline{p}_{n_k} - \frac{y^*_{n_k}}{2(1-\frac{y^*_{n_k}}{2(1-\overline{p}_{n_k})})} + t; \quad \forall \omega \in B_{n_k}; \ \forall k. \tag{43}$$

*By taking subsequences of* $(n_k)_{k=1}^{\infty}$ *(to make* $y^*_{n_k}$ *converge) we may assume w.l.o.g. that:*

$$\lim_{k\to\infty} \left( \overline{p}_{n_k} - \frac{y^*_{n_k}}{2(1-\frac{y^*_{n_k}}{2(1-\overline{p}_{n_k})})} + t \right) = \underline{p} - \frac{\underline{y}^* + \varepsilon}{2(1-\frac{\underline{y}^*+\varepsilon}{2(1-\underline{p})})} + t, \quad \text{for some } \ \varepsilon \geq 0.$$

*Thus, for some* $k_0$ *we have*

$$\overline{X}_{n_k}(\omega) < \underline{p} - \frac{\underline{y}^* + \varepsilon}{2(1-\frac{\underline{y}^*+\varepsilon}{2(1-\underline{p})})} + 2t; \ \forall \omega \in B_{n_k}; \ \forall k > k_0. \tag{44}$$

*Inserting* $\underline{y}^* = 2\underline{p}(1-\underline{p})$ *we have:*

$$\begin{aligned} \underline{p} - \frac{\underline{y}^* + \varepsilon}{2(1-\frac{\underline{y}^*+\varepsilon}{2(1-\underline{p})})} + 2t &\leq \underline{p} - \frac{\underline{y}^*}{2(1-\frac{\underline{y}^*}{2(1-\underline{p})})} + 2t \\ &= \underline{p} - \frac{2\underline{p}(1-\underline{p})}{2(1-\frac{2\underline{p}(1-\underline{p})}{2(1-\underline{p})})} + 2t = 2t, \end{aligned}$$

*implying that*

$$\overline{X}_{n_k}(\omega) < 2t; \ \forall \omega \in B_{n_k}; \ \forall k > k_0. \tag{45}$$

*As* $t > 0$ *is arbitrary, in particular, if* $2t < 1/2$*; since* $P(B_{n_k}) > q > 0$ *for all* $k$*, inequalities (45) imply that* $(X,P)$ *does not satisfy the* $CJT$*.*

*We conclude:* No distribution with $\frac{1}{2} \leq \underline{p} < 1$ and $\underline{y}^* = 2\underline{p}(1-\underline{p})$ satisfy the $CJT$.

Inspired by the previous example we move now to the proof of Theorem 22 stating the general necessary condition for the $CJT$ in $L_1$.

**Theorem 35.** *Let* $X = (X^1, X^2, \ldots, X^n, \ldots)$ *be sequence of binary random variables with joint distribution* $P$*. If* $\underline{y}^* > 2(2\underline{p}-1)(1-\underline{p})$*, then* $(X,P)$ *does not satisfy the* $CJT$*.*

*Proof.* Let $\tilde{x} = (2\underline{p}-1)(1-\underline{p})$ and notice that $x/(1-x/(1-p))$ is an increasing function for $x < 1-p$. Since $\underline{y}^*/2 > \tilde{x}$, let $t$ be such that

$$0 < t < \frac{1}{2}\left( \frac{\underline{y}^*}{2(1-\frac{\underline{y}^*}{2(1-\underline{p})})} - \frac{\tilde{x}}{1-\frac{\tilde{x}}{1-\underline{p}}} \right)$$

By Lemma 33, there exists a sequence of events $(B_{n_k})_{k=1}^{\infty}$ and $q > 0$, such that $P(B_{n_k}) > q > 0$ for all $k$, and (42) and, (by choosing an appropriate subsequence), (44) are satisfied. Thus, on these events we have,

$$\overline{X}_{n_k}(\omega) < \underline{p} - \frac{\underline{y}^* + \varepsilon}{2(1 - \frac{\underline{y}^*+\varepsilon}{2(1-\underline{p})})} + 2t \quad \leq \quad \underline{p} - \frac{\underline{y}^*}{2(1 - \frac{\underline{y}^*}{2(1-\underline{p})})} + 2t$$
$$< \quad \underline{p} - \frac{\tilde{x}}{1 - \frac{\tilde{x}}{1-\underline{p}}} + 2t - 2t.$$

Substituting $\tilde{x} = (2\underline{p} - 1)(1 - \underline{p})$ we have

$$\overline{X}_{n_k}(\omega) < \underline{p} - \frac{(2\underline{p} - 1)(1 - \underline{p})}{1 - (2\underline{p} - 1)} = \frac{1}{2}.$$

We conclude that $\overline{X}_{n_k}(\omega) < \frac{1}{2}$, for all $\omega \in B_{n_k}$ and for all $k > k_0$, implying that $(X, P)$ does not satisfy the $CJT$.

□

## 7.3 The $CJT$ and the Law of Large Numbers

At first sight, the asymptotic $CJT$ condition may look rather similar to the well-known *Law of Large Numbers* (*LLN*). It is the purpose of this section to clarify and state precisely the relationship between these two concepts.

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):

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P\left(|\overline{X}_n - \overline{p}_n| < \varepsilon\right) = 1 \tag{46}$$

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

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

Since by Proposition 20, 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 36.** *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 37.** *The statement of Proposition 36 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* $CJT$ *since it does not satisfy* $\lim_{n\to\infty} \sqrt{n}(\overline{p}_n - \frac{1}{2}) = \infty$ *which is a necessary and sufficient condition for* $CJT$ *(see Berend and Paroush (1998)).*

Unfortunately, Proposition 36 is of little use to us. This is due to the following fact:

**Proposition 38.** *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.

It follows thus from Proposition 38 that $LLN$ cannot hold when $\underline{y} > 0$ and thus we cannot use Proposition 36 to establish distributions in this region that satisfy the $CJT$.

Summing up, The $LLN$ and the $CJT$ are substantially two different properties that do not imply each other. The partial implication $LLN \Rightarrow CJT$ applies only for the horizontal line in $L_2$; $(\underline{p}, 0)$, for $\underline{p} > 1/2$, where the $CJT$ is easily established directly. Furthermore, all distributions with $\underline{y} > 0$ for which we established the validity of the $CJT$ do not satisfy the $LLN$.