The Annals of Probability
1979, Vol. 7, No. 4, 704–723

ON THE ZERO-ONE LAW FOR EXCHANGEABLE EVENTS

BY DAVID ALDOUS AND JIM PITMAN

University of Cambridge and University of California, Berkeley

We seek conditions for the exchangeable σ-field of an independent non-identically-distributed sequence of random variables to be trivial. A simple necessary condition is given, and this condition is shown to be sufficient when the range space is finite. In the case of a general range space, a stronger condition is shown to be sufficient.

1. Introduction. Let $X = (X_n)$ be a sequence of of independent random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{E} \subset \mathcal{F}$ be the exchangeable σ-field. That is, $\mathcal{E}$ is the collection of events of the form $\{X \in E\}$, where E is a measurable subset of sequence space which is invariant under the permutation of finitely many coordinates. Let μ_n be the distribution of X_n. If μ_n is the same for all n, then the well-known zero-one law of Hewitt and Savage (1955) asserts that $\mathcal{E}$ is trivial: $F \in \mathcal{E}$ implies $P(F) = 0$ or 1. The main purpose of this paper is to give some less restrictive conditions on (μ_n) which are sufficient for $\mathcal{E}$ to be trivial. These conditions are much weaker than those considered by Horn and Schach (1970), Blum and Pathak (1972) and Sendler (1975). Moreover, these conditions are necessary when the range space is finite.

Let the range of the random variables X_n be a measurable space $(S, \mathcal{S})$. So μ_n is the probability on $(S, \mathcal{S})$ defined by $\mu_n(B) = P(X_n \in B)$. Note that for a countable set S we always assume that $\mathcal{S}$ is the collection of all subsets of S, and write $\mu_n(x)$ instead of $\mu_n(\{x\})$, $x \in S$.

It appears that the way in which the sequence of distributions (μ_n) influences the structure of $\mathcal{E}$ becomes more complex as the cardinality of S increases. Thus, the results to be given take on different forms according to whether S is finite, countably infinite or uncountable.

Let us start with the simplest case of a two point set S.

(1.1) THEOREM. *Let $S = \{h, t\}$ and let $p_n = \mu_n(h)$, $q_n = \mu_n(t)$. A necessary and sufficient condition for $\mathcal{E}$ to be trivial is that $\sum p_n \wedge q_n = 0$ or ∞.*

REMARK. Here $\alpha \wedge \beta$ denotes the infimum of α and β. An equivalent condition is obtained upon replacing $p_n \wedge q_n$ by $p_n q_n$, because of the inequality

$$(1.2) \qquad \tfrac{1}{2}(\alpha \wedge \beta) \leqslant \alpha\beta/(\alpha + \beta) \leqslant \alpha \wedge \beta, \qquad \alpha, \beta \geqslant 0, \qquad \alpha + \beta > 0.$$

The sufficiency of this condition is a consequence of Theorem 1.6 below. We will prove the necessity now, because the proof is easy and the result is important for the formulation of conditions on (μ_n) in more general spaces. Let U_n be the

Received July 18, 1977; revised June 27, 1978.
AMS 1970 *subject classifications.* Primary 60F20; secondary 60J10.
Key words and phrases. Exchangeable events, zero-one law, Markov chain, coupling.

indicator of the event $\{X_n = h\}$, and let $V_n = U_n - p_n$. Then $EV_n = 0$ and $EV_n^2 = p_n q_n$. If $\Sigma p_n q_n < \infty$ then $Z = \Sigma V_n$ exists a.s. and satisfies $EZ = 0$, $EZ^2 = \Sigma p_n q_n$. Thus, in view of (1.2), if $0 < \Sigma p_n \wedge q_n < \infty$ then Z is an $\mathfrak{S}$-measurable random variable which is not a.s. constant, whence $\mathfrak{S}$ is not trivial.

It will be shown in Section 4 that when $\mathfrak{S}$ is not trivial, the above random variable Z takes only countably many values, and that $\mathfrak{S}$ is (up to null sets) the σ-field generated by Z. This sheds some light on a question raised at the end of the paper by Georgii (1976), which contains results concerning the exchangeable σ-field for dependent sequences of two valued random variables: see also Pitman (1978).

When S is the line, it is well known that $\mathfrak{S}$ includes the tail σ-field of the sequence of partial sums $(\Sigma_1^n X_i)$. It is easy to verify that the inclusion is an equality when S is a two-point subset of the line. This fact can be used to deduce the sufficiency in Theorem 1.1 from results of Orey (1966) concerning tail σ-fields (see also Mineka (1973)). In general, the tail σ-field of $(\Sigma_1^n X_i)$ may be trivial under conditions weaker than those given here for $\mathfrak{S}$ to be trivial. In this connection, recall a result of Lévy (1937): $P(\Sigma_1^n X_i > a_n$ infinitely often$) = 0$ or 1 for each sequence (a_n) of constants, provided that ΣX_i is essentially divergent.

The condition involved in the two-point case suggests consideration of the following condition for a general range space $(S, \mathbb{S})$.

CONDITION (a). For each $B \in \mathbb{S}$, $\Sigma \mu_n(B) \wedge \mu_n(B^c) = 0$ or ∞.

Given $B \in \mathbb{S}$, let 1_B denote its indicator function. Write $\mathfrak{S}_B$ for the exchangeable σ-field of the two-valued sequence $(1_B \circ X_n)$. Evidently

$$(1.3) \qquad \bigvee_{B \in \mathbb{S}} \mathfrak{S}_B \subset \mathfrak{S}.$$

Now the necessity in Theorem 1.1 gives a general necessary condition.

(1.4) COROLLARY. *Condition* (a) *is necessary for* $\mathfrak{S}$ *to be trivial.*

It turns out that Condition (a) is not sufficient when S is countably infinite (Example 7.3). So we are led to consider further conditions.

(1.5) DEFINITION. Let (μ_n) be a given sequence of probabilities on a countable S. For $x, y \in S$ call x and y *linked* if $\Sigma \mu_n(x) \wedge \mu_n(y) = \infty$. Call x and y *connected* if there exists a finite sequence of points, each linked to its successor, starting at x and ending at y.

CONDITION (b). For $x, y \in S$, x is connected to y whenever $\mu_n(x) \wedge \mu_n(y) > 0$ for some n.

(1.6) THEOREM. *Suppose S is finite. Then each of the conditions* (a) *and* (b) *is equivalent to the triviality of* $\mathfrak{S}$.

Proposition 2.7 will demonstrate that, for finite S, Conditions (a) and (b) are equivalent. Then the theorem is an immediate consequence of Corollary 1.4 and

Theorem 1.8 below. We note that it is not sufficient for condition (a) to hold only for singleton sets $B = \{x\}$, $x \in S$: see Example 7.1.

(1.7) THEOREM. *Suppose S is finite. Then $\mathfrak{G}$ coincides, up to null sets, with the σ-field generated by some countable partition of Ω.*

This theorem will be proved in Section 4. During the course of the proof it will be seen that, for finite S, there is essential equality in (1.3)—that is, $\mathfrak{G}$ coincides with $\bigvee_{B \in \mathbb{S}} \mathfrak{G}_B$ up to null sets. However, when S has more than two points these σ-fields are not identical (Example 7.2). Thus the sufficiency of condition (a) in the finite case does not seem to be an immediate consequence of the result (1.1) in the two-point case.

Let us now consider the case of countable S. It turns out that Theorem 1.6 does not completely extend to this case, but we have the following partial result, which will be proved in Section 3.

(1.8) THEOREM. *Suppose S is countable. Then Condition* (b) *is sufficient for $\mathfrak{G}$ to be trivial.*

This theorem and Corollary 1.4 yield the following implications, for countable S:

$$\text{Condition (b)} \Rightarrow \mathfrak{G} \text{ is trivial} \Rightarrow \text{Condition (a)}.$$

Theorem 1.6 asserts that both implications can be reversed, for finite S. But Examples 7.3 and 7.5 will show that neither implication can be reversed for infinite S.

Finally, let us consider the case of a general range space $(S, \mathbb{S})$. Obviously one would not expect Condition (b), involving measures of singletons, to be appropriate here.

(1.9) DEFINITION. For a given sequence (μ_n) of probabilities, let

$$\mathbb{S}_0 = \cap_n \{B \in \mathbb{S} : \mu_n(B) = 0 \text{ or } 1\}.$$

Observe that condition (a) is just the condition that $\sum \mu_n(B) \wedge \mu_n(B^c) = \infty$ for each $B \notin \mathbb{S}_0$. We now formulate a stronger condition, in terms of a reference measure.

CONDITION (c). There exists a probability measure ϕ on $(S, \mathbb{S})$ such that: for each $B \notin \mathbb{S}_0$ there exists $\varepsilon > 0$ such that

$$(1.10) \qquad \Sigma_n r_n(B, \varepsilon) = \infty, \qquad \text{where}$$

$$(1.11) \qquad r_n(B, \varepsilon) = \inf\{ \mu_n(B \setminus C) \wedge \mu_n(B^c \setminus C) : \phi(C) < \varepsilon\}.$$

If this condition holds for a given ϕ, we shall say (μ_n) is *ϕ-tame.*

(1.12) THEOREM. *Condition* (c) *is sufficient for $\mathfrak{G}$ to be trivial.*

For the proof see Section 6. If S is countable it is easy to see that conditions (b) and (c) are equivalent (Proposition 2.7), and so Theorem 1.8 is a consequence of Theorem 1.12. Nevertheless, we shall derive the results in the countable case independently of Theorem 1.12, because the proof of that theorem is both technical and complicated, and gives no insight into the structure of $\mathcal{E}$. The technique to be used in the countable case is much simpler and more intuitive, involving the novel application of some ideas from Markov chain theory.

The statement of Condition (c) is not as simple as one would wish. Sufficient conditions which may be easier to verify are given below. Their hypotheses will be shown to imply Condition (c) in Section 6. Let $\ll$ denote absolute continuity.

(1.13) COROLLARY. *Let Λ denote the set of limit points of (μ_n) in the total variation topology. Suppose that for each n there exists $\lambda_n \in \Lambda$ such that $\mu_n \ll \lambda_n$. Then $\mathcal{E}$ is trivial.*

(1.14) COROLLARY. *Let ϕ be a probability measure on $(S, \mathcal{S})$. Suppose*
 (i) *$\mu_n \ll \phi$, for each n;*
 (ii) *$\phi \ll \mu_n$ uniformly in n.*
Then $\mathcal{E}$ is trivial.

REMARKS. Condition (ii) means: for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu_n(B) \leqslant \delta$ implies $\phi(B) \leqslant \varepsilon$, $n \geqslant 1$. Corollary 1.13 is a generalisation of the result of Blum and Pathak (1972), who proved $\mathcal{E}$ trivial when $\mu_n \in \Lambda$ for each n. However, it is easy to construct examples where Λ is empty but the hypothesis of Corollary 1.14 holds.

In Section 2 we discuss the relations between the various conditions on (μ_n) which have been described. In Section 3 we develop a "coupling" technique, for countable S, and use it to prove Theorem 1.8. This technique is extended in Section 4 to describe the structure of $\mathcal{E}$ when it is nontrivial, and to prove Theorem 1.7. In Section 5 we digress to show how our results fit into the framework of Markov chain coupling theory. The case of a general $(S, \mathcal{S})$ is discussed in Section 6, where Theorem 1.12 is proved. Finally, Section 7 is a miscellany of examples.

We draw the reader's attention to the recent paper of Simons (1978), who discusses closure properties of the class of sequences for which $\mathcal{E}$ is trivial.

2. Relations between the conditions. Throughout this section S is countable and (μ_n) is a fixed sequence of probabilities on $(S, \mathcal{S})$, where $\mathcal{S}$ is the collection of all subsets of S. We shall be much concerned with equivalence relations on S, and the associated partitions of S into equivalence classes. Note that the relation $\sim$ is stronger than $\approx$ (that is, $x \sim y$ implies $x \approx y$) if and only if the associated partition $\mathfrak{R}^{\sim}$ is a refinement of $\mathfrak{R}^{\approx}$ (each element of $\mathfrak{R}^{\approx}$ is a union of elements of $\mathfrak{R}^{\sim}$). For $A, B \in \mathcal{S}$ write $\Sigma(A, B) = \Sigma\, \mu_n(A) \wedge \mu_n(B)$. Recall Definition 1.5 of *linked* and *connected*. In the present notation,

(2.1) $$x \text{ and } y \text{ are linked} \Leftrightarrow \Sigma(x, y) = \infty.$$

Clearly the relation of connection is an equivalence relation. Write $\mathcal{C}$ for the partition of equivalence classes. Then

(2.2) $$C \in \mathcal{C} \Rightarrow \Sigma(x, y) < \infty, \quad \text{each} \quad x \in C, \ y \in C^c.$$

Now let us say that x and y are *weakly linked* if $\Sigma(x, y) > 0$. Define *weakly* *connected* analogously, and write $w\mathcal{C}$ for the weak-connection partition. Since connection implies weak connection,

(2.3) $$\mathcal{C} \text{ is a refinement of } w\mathcal{C}.$$

Now recall Definition 1.9 of $\mathbb{S}_0$. In present notation $\mathbb{S}_0 = \{A \in \mathbb{S} : \Sigma(A, A^c) = 0\}$ and we clearly have

(2.4) $$B \notin \mathbb{S}_0 \Rightarrow \Sigma(x, y) > 0 \quad \text{for some } x \in B, y \in B^c,$$

(2.5) $$\mathbb{S}_0 \text{ is a } \sigma\text{-field, and } w\mathcal{C} \text{ is the set of atoms of } \mathbb{S}_0.$$

The main result of this section is Proposition 2.7 below. Here is a preliminary.

(2.6) LEMMA. *Let $B \subset S$. If*

$$\Sigma(x, x') < \infty, \quad \textit{for all} \quad x \in B, \quad x' \in B^c,$$

then

$$\Sigma(F, F') < \infty, \quad \textit{for all finite} \quad F \subset B, \quad F' \subset B^c.$$

PROOF. Apply repeatedly the inequality

$$(\beta + \gamma) \wedge \alpha \leqslant \beta \wedge \alpha + \gamma \wedge \alpha; \qquad \alpha, \beta, \gamma \geqslant 0.$$

(2.7) PROPOSITION. *The following are equivalent.*

(a′) *For each $B \notin \mathbb{S}_0$, there exist finite sets $F \subset B$, $F' \subset B^c$ such that $\Sigma(F, F') = \infty$.*

(a″) *For each $B \notin \mathbb{S}_0$, there exist $x \in B, y \in B^c$ such that $\Sigma(x, y) = \infty$.*

(b′) $\mathcal{C} \subset \mathbb{S}_0$.

(b) *For $x, y \in S$, $\Sigma(x, y) > 0$ implies that x is connected to y.*

(c) *The sequence (μ_n) is ϕ-tame, for some probability ϕ on* $(S, \mathbb{S})$.

REMARK. Here (b) and (c) are just Conditions (b) and (c) of Section 1. Moreover (a′) is the same as Condition (a) of Section 1 *when S is finite*. Hence Proposition 2.7 enables Theorem 1.6 to be deduced from Corollary 1.4 and Theorem 1.8. The proof of Proposition 2.7 is achieved by proving the circle of implications $(a') \Rightarrow (b') \Rightarrow (b) \Rightarrow (a'') \Rightarrow (c) \Rightarrow (a')$. Observe, however, that since the implication $(a'') \Rightarrow (a')$ is obvious the notion of ϕ-tameness is not needed for the equivalence of the other conditions.

PROOF. $(a') \Rightarrow (b')$. Consider $C \in \mathcal{C}$. By (2.2) and Lemma 2.6 we see that $\Sigma(F, F') < \infty$ for each $F \subset C$, $F' \subset C^c$. Thus (a') implies $C \in \mathbb{S}_0$.

$(b') \Rightarrow (b)$. If (b') holds then (2.5) implies that $w\mathcal{C}$ is a refinement of $\mathcal{C}$. Thus by (2.3), $w\mathcal{C} = \mathcal{C}$, and the relations of connection and weak connection are therefore identical.

$(b) \Rightarrow (a'')$. Consider $B \notin \mathcal{S}_0$. By (2.4) and (b) there exists a point $x \in B$ which is connected to some point $y \in B^c$. Hence some points $x' \in B, y' \in B^c$ are linked. Now use (2.1).

$(a'') \Rightarrow (c)$. Let ϕ be a probability on $(S, \mathcal{S})$ such that $\phi(s) > 0$ for each $s \in S$. Consider $B \notin \mathcal{S}_0$, and let x, y be as in (a''). Choose $0 < \varepsilon < \phi(x) \wedge \phi(y)$. Recalling Definition 1.11, we see that $r_n(B, \varepsilon) \geqslant \mu_n(x) \wedge \mu_n(y)$. Hence $\Sigma_n r_n \ (B, \varepsilon) \geqslant \Sigma(x, y) = \infty$, and (μ_n) is ϕ-tame.

$(c) \Rightarrow (a')$. Suppose (μ_n) is ϕ-tame for some ϕ. Consider $B \notin \mathcal{S}_0$. Choose $\varepsilon > 0$ such that $\Sigma_n r_n(B, \varepsilon) = \infty$. Now choose a finite set F_0 such that $\phi(F_0) > 1 - \varepsilon$, and let $F = F_0 \cap B$, $F' = F_0 \cap B^c$. The definition of r_n shows that $r_n(B, \varepsilon) \leqslant \mu_n(F) \wedge \mu_n(F')$. Hence $\Sigma(F, F') = \infty$, and (a') holds.

We end this section by proving, for use in Section 4, a partial analogue of (2.5) for the connection partition.

(2.8) DEFINITION. Let $\mathcal{S}_+ = \{B \subset S : \Sigma(B, B^c) < \infty\}$.

(2.9) PROPOSITION. *Suppose S is finite. Then $\mathcal{S}_+$ is the field generated by the connection partition $\mathcal{C}$.*

PROOF. Consider $B \in \mathcal{S}_+$. Suppose $x \in B$ and x is linked to y. Then $y \in B$, because $\Sigma(B, B^c) < \infty$. Hence B is a union of sets in $\mathcal{C}$. Conversely, suppose B is a union of sets in $\mathcal{C}$. Then $\Sigma(x, y) < \infty$ for each $x \in B, y \in B^c$. Now Lemma 2.6 and the finiteness of S imply $\Sigma(B, B^c) < \infty$, i.e., $B \in \mathcal{S}_+$.

REMARK. If S is countable it is easy to see that $\mathcal{S}_+$ is a field contained in the σ-field generated by $\mathcal{C}$. But $\mathcal{S}_+$ is not necessarily a σ-field: see Example 7.4.

3. The coupling argument. We begin with an informal outline of the proof of Theorem 1.8, which exploits a particular kind of coupling of processes. Let a sequence (X_n) of independent random variables and points y, z in S be given. Suppose that on some probability space $(\Omega, \mathcal{F}, P)$ one can define copies (Y_n) and (Z_n) of (X_n) such that for each $\omega \in \Omega$ the sequence $(y, Y_2(\omega), Y_3(\omega), \cdots)$ is a finite permutation of $(z, Z_2(\omega), Z_3(\omega), \cdots)$. Suppose further that for each $m \geqslant 1$ this can also be done with $(X_n, n \geqslant 1)$ replaced by $(X_{m+n}, n \geqslant 1)$. We shall then say that y and z can be *coupled*. The proof of Theorem 1.8 divides into two parts. The first (Proposition 3.13) shows that the exchangeable σ-field is trivial provided that each weakly linked pair of points can be coupled. The second (Proposition 3.14) gives a construction to show that linked points can be coupled. From these facts and a transitivity argument it follows that $\mathcal{E}$ is trivial provided that weakly linked points are connected: this is Theorem 1.8.

The coupling construction involved here can be represented as a coupling of Markov chains of the kind described by Griffeath (1975). This point of view will be developed in Section 5, but it is not adopted here.

It is convenient to set up the formal machinery on sequence space. Let S be a countable set. Write S^∞ for the space of sequences $\mathbf{x} = (x_i)$ of elements of S. Write

$X_n : S^\infty \to S$ for the coordinate map $X_n(\mathbf{x}) = x_n$, and let $\mathbb{S}^\infty$ be the σ-field on S^∞ generated by the coordinate maps. Let G be the group of finite permutations of $N = \{1, 2, \ldots\}$, and note that G is countable. We regard $\sigma \in G$ as acting on $\mathbf{x} \in S^\infty$:$(\sigma\mathbf{x})_i = x_{\sigma(i)}$. Then the exchangeable σ-field $\mathcal{E}$ is defined as the collection of sets $E \in \mathbb{S}^\infty$ such that $\mathbf{x} \in E$, $\sigma \in G \Rightarrow \sigma\mathbf{x} \in E$. Now let Q be an arbitrary probability measure on $(S^\infty, \mathbb{S}^\infty)$. Under Q, (X_n) forms a sequence of S-valued random variables, which may be dependent.

(3.1) DEFINITION. Let $y, z \in S$, $m \in N$. A *Q post-m coupling which exchanges y and z*, or more briefly a *(Q, m, y, z) coupling* is a pair $(\mathbf{Y}, \mathbf{Z})$ of sequences of random variables $\mathbf{Y} = (Y_i)$, $\mathbf{Z} = (Z_i)$ defined on a common probability space $(\Omega, \mathcal{F}, P)$ and such that:

(3.2) $$Y_1 = y, \qquad Z_1 = z;$$

(3.3) $$Y_i = Z_i, \qquad 2 \leqslant i \leqslant m$$

(3.4) $(Y_i, i > m)$and $(Z_i, i > m)$ have the same distribution as $(X_i, i > m)$ under Q;

(3.5) there exists a random permutation $\sigma:\Omega \to G$ such that $\mathbf{Y} = \sigma\mathbf{Z}$.

(3.6) DEFINITION. Let Q be given. For $y, z \in S$ say *y and z can be coupled,* and write $y \sim z$, if for each m there exists a (Q, m, y, z) coupling. Call $\sim$ the *coupling relation.*

(3.7) PROPOSITION. *The coupling relation is an equivalence relation.*

PROOF. Only transitivity needs verifying. Suppose there exists a (Q, m, w, z) coupling $(\mathbf{W}, \mathbf{Y})$ and a (Q, m, y, z) coupling $(\mathbf{Y}', \mathbf{Z}')$. Let $s \in S$ be arbitrary. We may clearly assume that $W_i = Y_i = Y_i' = Z_i' = s$ for $2 \leqslant i \leqslant m$. Let $\mathbf{Y}^*$ be a sequence with the common distribution of $\mathbf{Y}$ and $\mathbf{Y}'$; then let $\mathbf{W}^*$ and $\mathbf{Z}^*$ be conditionally independent given $\mathbf{Y}^*$, with the conditional distribution of $\mathbf{W}^*$ given $\mathbf{Y}^*$ equal to that of $\mathbf{W}$ given $\mathbf{Y}$, and the conditional distribution of $\mathbf{Z}^*$ given $\mathbf{Y}^*$ equal to that of $\mathbf{Z}'$ given $\mathbf{Y}'$. Then there exist random permutations σ and π such that $\mathbf{W}^* = \sigma\mathbf{Y}^*$ a.s. and $\mathbf{Y}^* = \pi\mathbf{Z}^*$ a.s. After discarding the null sets, one sees that $\mathbf{W}^* = (\sigma\pi)\mathbf{Z}^*$ and hence that $(\mathbf{W}^*, \mathbf{Z}^*)$ is a (Q, m, w, z) coupling.

(3.8) REMARK. If there exists a (Q, m, y, z) coupling, there also exists a (Q, n, y, z) coupling for each $n < m$.

(3.9) REMARK. If $y \neq z$, there cannot exist $k < \infty$ such that each random permutation $\sigma(\omega)$ moves only the first k co-ordinates. For this would imply $1 + \sum_{i=2}^k 1(Y_i = y) = \sum_{i=2}^k 1(Z_i = y)$, which is impossible because $P(Y_i = y) = P(Z_i = y)$ for $i = 2, \cdots, k$.

(3.10) DEFINITION. Given a probability Q on $(S^\infty, \mathbb{S}^\infty)$, points $x_1, \cdots, x_n \in S$

and $A \in \mathcal{S}^\infty$, let

$$Q(x_1, \cdots, x_n, A) = Q\{(x_1, \cdots, x_n, X_{n+1}, X_{n+2}, \cdots) \in A\}.$$

(3.11) REMARK. If $E \in \mathcal{E}$ then $Q(x_1, \cdots, x_n, E)$ is a symmetric function of $x_1, \cdots, x_n$.

(3.12) LEMMA. *Suppose $E \in \mathcal{E}$ and $z_i \sim y_i$, $1 \leq i \leq n$. Then $Q(y_1, \cdots, y_n, E) = Q(z_1, \cdots, z_n, E)$.*

PROOF. Because (z_i) may be changed to (y_i) by altering one term at a time, we may assume that (z_i) and (y_i) differ only at a single index. And by the symmetry of $Q(-, E)$ we may assume this index to be 1. So suppose $y_1 \neq z_1$, $y_i = z_i$ for $2 \leq i \leq n$. Let $(\mathbf{Y}, \mathbf{Z})$ be a (Q, n, y_1, z_1) coupling. By Definition 3.1 we may assume that $Y_i = Z_i = y_i$, $2 \leq i \leq n$. Now by (3.4)

$$Q(y_1, \cdots, y_n, E) = P(\mathbf{Y} \in E)$$
$$Q(z_1, \cdots, z_n, E) = P(\mathbf{Z} \in E).$$

But the events $\{\mathbf{Y} \in E\}$ and $\{\mathbf{Z} \in E\}$ are identical, because $\mathbf{Y}(\omega) = \sigma(\omega)\mathbf{Z}(\omega)$.

The preceding results are valid for an arbitrary probability Q. But now we need to assume that Q is a product measure, i.e., that (X_n) is an independent sequence of random variables. So let (μ_n) be a sequence of probability measures on S, and recall the definition of *weakly linked* from Section 2.

(3.13) PROPOSITION. *Suppose $Q = \Pi_n \mu_n$. If y and z can be coupled whenever y and z are weakly linked, then $\mathcal{E}$ is trivial under Q.*

REMARK. The converse also holds: see Proposition 5.5.

PROOF. Fix $E \in \mathcal{E}$. Write $\mathcal{F}_n$ for the σ-field generated by $(X_1, \cdots, X_n)$. By independence, $Q(E|\mathcal{F}_n) = Q(X_1, \cdots, X_n, E)$. Now let $(y_1, \cdots, y_n)$ and $(z_1, \cdots, z_n)$ be two sequences, each of which is a sequence of possible values of $(X_1, \cdots, X_n)$: that is, $\mu_i(y_i) > 0$ and $\mu_i(z_i) > 0$, $1 \leq i \leq n$. Then y_i and z_i are weakly linked; by hypothesis $y_i \sim z_i$. Using Lemma 3.12, we see that $Q(E|\mathcal{F}_n)$ is a.s. constant for each n. But $Q(E|\mathcal{F}_n) \to 1_E$ a.s., whence $Q(E) = 0$ or 1.

This completes the first part of the argument. We now state the result of the second part, but before embarking on the proof we will show how to deduce Theorem 1.8.

(3.14) PROPOSITION. *Suppose $Q = \Pi_n \mu_n$. If y is linked to z there exists a $(Q, 1, y, z)$ coupling.*

PROOF OF THEOREM 1.8. Let (μ_n) be a sequence of probabilities satisfying Condition (b). Suppose y and z are linked under $Q = \Pi_n \mu_n$. Fix $m \geq 1$, and let $\mu_n' = \mu_{m+n}$. Then y and z are linked under $Q' = \Pi_n \mu_n'$ also. Applying Proposition

3.14 to Q', there exists a $(Q', 1, y, z)$ coupling $(\mathbf{Y}, \mathbf{Z})$. Now define

$$\begin{aligned}(Y_1', Z_1') &= (y, z);\\ (Y_i', Z_i') &= (s, s); \qquad 2 \leqslant i \leqslant m;\\ &= (Y_{i-m+1}, Z_{i-m+1}), \qquad i > m;\end{aligned}$$

where $s \in S$ is arbitrary. Then (Y', Z') is a (Q, m, y, z) coupling. Thus y and z can be coupled under $Q = \Pi \mu_n$. We have proved this for any linked pair y, z but the coupling relation is an equivalence relation, so

(3.15) y and z can be coupled whenever y and z are connected.

Now Condition (b) asserts that y and z are connected whenever y and z are weakly linked. Thus Condition (b) implies the hypothesis of Proposition 3.13, and hence the triviality of $\mathcal{E}$.

(3.16) DEFINITION. Given $\mathbf{x} = (x_i) \in S^\infty, y \in S$, let

$$M_n(\mathbf{x}, y) = \Sigma_{i=1}^n 1(x_i = y).$$

Let $M_n(\mathbf{x})$ denote the vector $(M_n(\mathbf{x}, y); y \in S)$. Then M_n, considered as a random vector defined on $(S^\infty, \mathcal{S}^\infty, Q)$, describes the empirical distribution of $(X_1, \cdots, X_n)$ under Q.

The next lemma is an immediate consequence of the definition.

(3.17) LEMMA. *$M_n(\mathbf{x}) = M_n(\mathbf{y})$ if and only if there exists a permutation σ of $\{1, \cdots, n\}$ such that $x_i = y_{\sigma(i)}$, $1 \leqslant i \leqslant n$.*

PROOF OF PROPOSITION 3.14. Let y and z be linked. Let $Y_1 = y$, $Z_1 = z$. Let $((Y_n, Z_n), n \geqslant 2)$ be a sequence of independent $S \times S$ valued random variables, whose laws are specified as follows:

(i) Y_n has law μ_n;

(ii) if $Y_n \notin \{y, z\}$ then $Z_n = Y_n$;

(iii) conditional on $\{Y_n = y\}$ or $\{Y_n = z\}$ the law of Z_n is μ_n conditioned on $\{y, z\}$.

It is clear that both $(Y_n, n \geqslant 2)$ and $(Z_n, n \geqslant 2)$ have law $\Pi_2^\infty \mu_n$.

Notice that $M_n(\mathbf{Y}, s) = M_n(\mathbf{Z}, s)$ for all $s \notin \{y, z\}$. Hence $M_n(\mathbf{Y}) = M_n(\mathbf{Z})$ iff $M_n(\mathbf{Y}, y) = M_n(\mathbf{Z}, y)$. Define

$$\begin{aligned}D_n &= M_n(\mathbf{Y}, y) - M_n(\mathbf{Z}, y),\\ T &= \inf\{n : D_n = 0\},\end{aligned}$$

and suppose we can prove

(3.18) $$P(T < \infty) = 1.$$

Then we can define

(3.19) $$\begin{aligned}Z_n^*(\omega) &= Z_n(\omega) \quad \text{if} \quad n \leqslant T(\omega) < \infty\\ &= Y_n(\omega) \quad \text{otherwise.}\end{aligned}$$

We claim that the sequence $(Z_n^*, n \geq 2)$ has the same law as $(Z_n, n \geq 2)$. It clearly suffices to verify this conditional on $\{T = m\}$, for each $m \geq 2$. But conditional on $\{T = m\}$, each of the sequences $(Y_n, n > m)$ and $(Z_n, n > m)$ is independent of $(Z_2, \cdots, Z_m)$ with law $\Pi_{n>m}\mu_n$, and we have Z_n^* equal to $Z_n(2 \leq n \leq m)$ and equal to $Y_n(n > m)$.

Clearly

$$M_T(\mathbf{Z}^*) = M_T(\mathbf{Z}) = M_T(\mathbf{Y}) \qquad \text{on } \{T < \infty\}$$

$$Z_n^* = Y_n \qquad \text{on } \{T < n\}.$$

So by Lemma 3.17 there exists a finite permutation $\sigma(\omega)$ such that $\mathbf{Y}(\omega) = \sigma(\omega)\mathbf{Z}^*(\omega)$. Thus $(\mathbf{Y}, \mathbf{Z}^*)$ is a $(Q, 1, y, z)$ coupling. To prove (3.18) observe that (D_n) is a random walk on the integers with nonstationary jump probabilities

$$P(D_n = D_{n-1} \pm 1) = \alpha_n$$

$$P(D_n = D_{n-1}) = 1 - 2\alpha_n,$$

where $\alpha_n = \mu_n(y)\mu_n(z)/(\mu_n(y) + \mu_n(z))$. But y and z are linked, so the inequality (1.2) implies $\Sigma\alpha_n = \infty$. Thus (D_n) jumps infinitely often a.s. by the Borel-Cantelli lemma. But (D_n) watched only when it jumps is a simple symmetric random walk. This motion hits zero a.s., hence so too does (D_n).

4. The structure of nontrivial $\mathcal{E}$. We continue working in the framework set up in the last section, but now aim to describe the structure of $\mathcal{E}$, when it is not trivial. We start with some definitions and simple lemmas. No measure is involved yet.

Let $\mathcal{R}$ be a partition of S. Recall the definition (3.16) of $M_n(\mathbf{x}, y)$.

(4.1) DEFINITION. For $\mathbf{x} \in S^\infty, A \subset S$ let $M_n(\mathbf{x}, A) = \Sigma_{y \in A} M_n(\mathbf{x}, y)$. Let $M_n^{\mathcal{R}}(\mathbf{x})$ be the vector $(M_n(\mathbf{x}, A), A \in \mathcal{R})$.

(4.2) DEFINITION. For $\mathbf{x} \in S$ let $X_n^{\mathcal{R}}(\mathbf{x})$ be the element $R \in \mathcal{R}$ which contains $X_n(\mathbf{x})$.

Thus $(X_1^{\mathcal{R}}, X_2^{\mathcal{R}}, \cdots)$ describes the coordinate sequence $(X_1, X_2, \cdots)$ modulo $\mathcal{R}$, and $M_n^{\mathcal{R}}$ is the empirical distribution of $X_1^{\mathcal{R}}, \cdots, X_n^{\mathcal{R}}$.

(4.3) NOTATION. Let $\mathcal{E}(Y_n; n \geq 1)$ denote the exchangeable σ-field of a sequence $(Y_n; n \geq 1)$, and $\mathcal{T}(Y_n; n \geq 1)$ the tail σ-field.

The first two lemmas below are obvious.

(4.4) LEMMA. *If $\mathcal{P}$ is a refinement of $\mathcal{R}$ then $\mathcal{E}(X_n^{\mathcal{R}}; n \geq 1) \subset \mathcal{E}(X_n^{\mathcal{P}}; n \geq 1)$. In particular $\mathcal{E}(X_n^{\mathcal{R}}; n \geq 1) \subset \mathcal{E}$.*

(4.5) LEMMA. *$M_n^{\mathcal{R}}(\mathbf{x}) = M_n^{\mathcal{R}}(\mathbf{y})$ if and only if there exists a permutation σ of $\{1, \cdots, n\}$ such that, for each i, the pair $x_i, y_{\sigma(i)}$ lie in the same element of $\mathcal{R}$.*

(4.6) LEMMA. $\mathcal{E}(X_n^{\mathcal{R}}; n \geq 1) = \mathcal{T}(M_n^{\mathcal{R}}; n \geq 1)$.

PROOF. Let $\mathcal{G}_n$ be the σ-field generated by $(M_n^{\mathfrak{R}}, M_{n+1}^{\mathfrak{R}}, \cdots)$. Then $\mathcal{G}_n$ is also generated by $(M_n^{\mathfrak{R}}, X_{n+1}^{\mathfrak{R}}, X_{n+2}^{\mathfrak{R}}, \cdots)$. It then follows from Lemma 4.5 that $\mathcal{G}_n$ is the collection of sets of the form $\{(X_i^{\mathfrak{R}}) \in E_n\}$, where $E_n \subset S^\infty$ is a measurable set invariant under permutation of coordinates $\{1, \cdots, n\}$. Hence $\mathcal{E}(X_n^{\mathfrak{R}}; n \geq 1) = \cap \mathcal{G}_n = \mathcal{T}(M_n^{\mathfrak{R}}; n \geq 1)$.

We now introduce a product law $Q = \prod \mu_n$ on S^∞. Let $\mathfrak{R}$ be the partition generated by the coupling relation (3.6). Let $\mathcal{C}$ be the partition generated by the connection relation (1.5). Given a σ-field $\mathfrak{D} \subset \mathcal{E}$, let us say that $\mathfrak{D} = \mathcal{E}$ a.s. if for each set $E \in \mathcal{E}$ there exists $D \in \mathfrak{D}$ such that $Q(D \setminus E) = Q(E \setminus D) = 0$.

(4.7) PROPOSITION. $\mathcal{E}(X_n^{\mathfrak{R}}; n \geq 1) = \mathcal{E}(X_n^{\mathcal{C}}; n \geq 1) = \mathcal{E}$ *a.s.*

PROOF. By (3.15) $\mathcal{C}$ is a refinement of $\mathfrak{R}$. So by Lemma 4.4, $\mathcal{E}(X_n^{\mathfrak{R}}; n \geq 1) \subset \mathcal{E}(X_n^{\mathcal{C}} : n \geq 1) \subset \mathcal{E}$. Now fix $E \in \mathcal{E}$. Using Lemmas 4.5 and 3.12 as in the proof of Proposition 3.13, we see that $Q(E|\mathcal{F}_n)$ is a function of $M_n^{\mathfrak{R}}$. Hence E is a.s. equal to some event in $\mathcal{T}(M_n^{\mathfrak{R}}; n \geq 1)$, and the result follows from Lemma 4.6.

(4.8) PROPOSITION. *Suppose S is finite. Then there exists a countably-valued random variable Z such that $\sigma(Z) = \mathcal{E}$* a.s.

PROOF. By Proposition 2.9,

$$\Sigma \mu_n(C) \wedge \mu_n(C^c) < \infty \qquad \text{each} \quad C \in \mathcal{C}. \tag{4.9}$$

Now $\mathcal{C}$ has at most $|S|$ elements. Thus we may choose n_0 so large that

$$\mu_n(C) \wedge \mu_n(C^c) < 1/|S|, \qquad C \in \mathcal{C}, \qquad n \geq n_0.$$

Then there exist elements C_n, $n \geq n_0$, of $\mathcal{C}$ such that

$$\mu_n(C_n) > \tfrac{1}{2}, \tag{4.10}$$

for otherwise $\mu_n(C) < 1/|S|$ for each $C \in \mathcal{C}$, which contradicts $\Sigma_C \mu_n(C) = 1$. Choose $\mathbf{x} \in S^\infty$ such that $x_n \in C_n$, $n \geq n_0$. Let $m_n^{\mathcal{C}}$ be the vector $M_n^{\mathcal{C}}(\mathbf{x})$. Consider the random vector $Z_n = M_n^{\mathcal{C}} - m_n^{\mathcal{C}}$. For each $K \geq n_0$,

$$\begin{aligned} Q(Z_n \neq Z_K \quad \text{for some} \quad n > K) &= Q(X_n \notin C_n \quad \text{for some} \quad n > K) \\ &\leq \Sigma_{n>K} \mu_n(C_n^c) \\ &= \Sigma_{n>K} \mu_n(C_n) \wedge \mu_n(C_n^c) \quad \text{by} \quad (4.10) \\ &\leq \Sigma_{C \in \mathcal{C}} \Sigma_{n>K} \mu_n(C) \wedge \mu_n(C^c) \end{aligned}$$

which converges to zero as $K \to \infty$, by (4.9). Thus there exist an a.s. finite random variable N and a random vector Z such that

$$Z_n = Z, \qquad n \geq N. \tag{4.11}$$

Now Z has a finite number of integer components, and so is countably valued.

And

$$\mathfrak{T}(M_n^{\mathcal{C}}; n \geqslant 1) = \mathfrak{T}(Z_n; n \geqslant 1)$$
$$= \sigma(Z) \quad \text{a.s. by} \quad (4.11).$$

The result now follows from Lemma 4.6 and Proposition 4.7.

5. The associated Markov chain. In this section we show how the empirical distribution process may be regarded as a Markov chain. We show that Proposition 3.13 (if weakly linked points can be coupled then $\mathcal{E}$ is trivial) is a special case of a result concerning Markov chain coupling, and prove its converse.

We begin with a general discussion of Markov chains. Let Λ be a countable set, and let $p = (p(\kappa, \lambda) : \kappa, \lambda \in \Lambda)$ be a Markov matrix. A Λ-valued process L is *Markov* (p) if L is a Markov chain with discrete time set $\{0, 1, 2, \cdots\}$ and stationary transition probabilities p. A bivariate process $(K, L) = ((K_n, L_n), n \geqslant 0)$ defined on a probability space $(\Omega, \mathfrak{F}, P)$ is a (κ, λ, p)*-coupling* if

(a) $K_0 = \kappa$ and K is Markov (p);

(b) $L_0 = \lambda$ and L is Markov (p);

(c) $K_n(\omega) = L_n(\omega)$ for $n \geqslant T(\omega)$,

where $T : \Omega \to \{0, 1, \cdots\}$ is an a.s. finite random time. Two points κ, λ in Λ can be *p-coupled* if there exists a (κ, λ, p)-coupling. According to Theorem 4 of Griffeath (1975),

(5.1) κ and λ can be p-coupled iff $\Delta(\kappa, \lambda) = 0$, where

$$\Delta(\kappa, \lambda) = \lim_{n\to\infty} \| p^n(\kappa, \cdot) - p^n(\lambda, \cdot) \|$$

and $\| \cdot \|$ is the total variation norm (see also Pitman (1976)). Now suppose M is a process, and for each $\gamma \in \Lambda$ let P_γ be a probability under which $M_0 = \gamma$ a.s. and M is Markov (p). From (5.1) and Proposition 1 of Griffeath (1975), the following assertions are equivalent.

(a) All pairs of states can be p-coupled.

(b) For each A in the tail σ-field $\mathfrak{T}(M)$, either $P_\gamma(A) = 0$ for all $\gamma \in \Lambda$ or $P_\gamma(A) = 1$ for all $\gamma \in \Lambda$.

However, we require a slightly different form of this result which applies when there is a given initial distribution, and where $\mathfrak{T}(M)$ may be trivial for this initial distribution but not for others.

(5.2) PROPOSITION. *Let M be Markov (p), and let q_n be the distribution of M_n. The following statements are equivalent.*

(a) *The tail σ-field $\mathfrak{T}(M)$ is trivial.*

(b) *Two states κ and λ can be p-coupled whenever $q_n(\kappa) \wedge q_n(\lambda) > 0$ for some n.*

PROOF. By virtually the same argument as that used to prove (3.12), if κ can be p-coupled to λ then for $A \in \mathfrak{T}(M)$, $P(A \mid M_n = \kappa) = P(A \mid M_n = \lambda)$. Then if (b) holds $P(A \mid M_1, \cdots, M_n)$ is a.s. constant for each n and (a) follows by martingale convergence.

Suppose now that (a) holds. Let $\mathfrak{T}_j$ be the σ-field generated by $(M_j, M_{j+1}, \cdots)$. Define

$$\varepsilon(j, G) = \sup_{F \in \mathfrak{T}_j} |P(F|G) - P(F)|; \qquad G \in \mathfrak{T}_0, \qquad P(G) > 0.$$

According to (1.129) of Freedman (1971), triviality of $\mathfrak{T}(M)$ implies $\lim_j \varepsilon(j, G) = 0$ for each G. By considering $G = \{M_n = \kappa\}$, where $P(G) = q_n(\kappa) > 0$, and F of the form $F = \{M_{n+j} \in B\}$, $B \subset \Lambda$, we see that

$$\|p^j(\kappa, \cdot) - q_{n+j}(\cdot)\| \leqslant \varepsilon(j, G).$$

So if $q_n(\kappa) \wedge q_n(\lambda) > 0$ we deduce that $\Delta(\kappa, \lambda) = 0$ and then (b) results from (5.1)

Now as in Section 3 let (M_n) be the empirical distribution process for the coordinate sequence (X_n) on S^∞, and recall from Lemma 4.6 that the exchangeable σ-field $\mathfrak{E}$ is identical to $\mathfrak{T}(M)$.

(5.3) PROPOSITION. *Under a product measure $Q = \Pi\,\mu_n$ the process (M_n) is a Markov chain with stationary transition probabilities.*

PROOF. Let Λ be the set of finite counting measures, i.e., functions $\lambda : S \to \{0, 1, 2, \cdots\}$ such that $|\lambda| = \Sigma\lambda(x) < \infty$. Note that Λ is countable. The process (M_n) takes values in Λ, and is easily verified to be Markov with the following transition matrix $\{p(\kappa, \lambda)\}$:

$$\begin{aligned} p(\kappa, \lambda) &= \mu_{n+1}(x) \quad \text{if} \quad |\kappa| = n \quad \text{and} \quad \lambda = \kappa + \delta_x, \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

where δ_x is the unit mass at $x \in S$.

We keep the above notation for the remainder of the section.

(5.4) LEMMA. *Let $\kappa, \lambda \in \Lambda$ and $y, z \in S$, and suppose that $|\kappa| = |\lambda| = m$ and $\kappa - \lambda = \delta_y - \delta_z$. The counting measures κ and λ can be p-coupled if and only if there exists a Q-post-m coupling which exchanges the points y and z.*

REMARK. The "if" part of the lemma is obvious. From this and Proposition 5.2 we deduce Proposition 3.13, using the transitivity of the p-coupling relation (cf. 3.6).

PROOF. Let (K, L) be a (κ, λ, p) coupling defined on $(\Omega, \mathfrak{F}, P)$, but index (K, L) by $\{m, m + 1, \cdots\}$ instead of $\{0, 1, \cdots\}$. By discarding a null set we can assume that for all $n \geqslant m$ there exist S-valued random variables Y_n and Z_n such that for all $\omega \in \Omega$, $n \geqslant m$ the following is true: $K_{n+1}(\omega)$ equals $K_n(\omega)$ plus a unit mass at $Y_n(\omega)$, and $L_{n+1}(\omega)$ equals $L_n(\omega)$ plus a unit mass at $Z_n(\omega)$. Define also $Y_1 = y$, $Z_1 = z$, $Y_i = Z_i = y_i$, $2 \leqslant i \leqslant m$, where $(y_i, 2 \leqslant i \leqslant m)$ is a sequence with empirical distribution $\kappa - \delta_y = \lambda - \delta_z$. Then (Y, Z) is evidently a (Q, m, y, z) coupling.

Now we can prove the converse to Proposition 3.13.

(5.5) PROPOSITION. *Suppose $\mathfrak{E}$ is trivial. Then y and z can be coupled whenever y and z are weakly linked.*

PROOF. Suppose $\mathcal{E}$ is trivial, and fix $y \neq z$, $m \geqslant 1$ such that $\mu_m(y) \wedge \mu_m(z) > 0$. Let $\mathbf{y}$ be a sequence with $y_m = y$ and $\mu_n(y_n) > 0$ for all n, and let $z_m = z$, $z_n = y_n$ for $n \neq m$. Let $\kappa_n = M_n(\mathbf{y})$, $\lambda_n = M_n(\mathbf{z})$. Then $\kappa_n - \lambda_n = \delta_y - \delta_z$, $n \geqslant m$, and $Q(M_n = \lambda) > 0$ for both $\lambda = \lambda_n$ and $\lambda = \kappa_n$. Since $\mathcal{E} = \mathcal{T}(M)$, κ_n can be p-coupled to λ_n by (5.2). Thus we see by (5.4) that for each $n \geqslant m$ there exists a (Q, n, y, z) coupling, i.e., y is coupling equivalent to z.

REMARK. It is now clear that the following are equivalent:

(i) $\mathcal{E}$ is trivial,

(ii) y is coupling equivalent to z whenever y and z are weakly linked,

(iii) the coupling relation is identical to the relation of weak connection,

(iv) the partition generated by the coupling relation is identical to the collection of atoms of the σ-field $\mathcal{S}_0$.

Recall Theorem 1.8 asserted that $\mathcal{E}$ is trivial provided weakly linked points are linked. This was proved via Proposition 3.14, which showed that linked points can be coupled. Now the converse to Proposition 3.14 is false (Example 7.5). So any method for constructing couplings under hypotheses weaker than those of Proposition 3.14 would lead to an improvement of Theorem 1.8. We do not know of any general method, but in Example 7.5 an ad hoc coupling is constructed where no pair of distinct points is linked.

A version of Proposition 3.14 can be proved for more general (i.e., uncountable) range spaces. So it would be natural to try to prove the zero-one law on more general spaces using this same technique. However, on more general spaces the problem of constructing couplings using hypotheses on (μ_n) seems very difficult, since obviously one cannot merely watch singletons. We are therefore forced to use different methods.

6. The general range space. Here we consider independent random variables with values in an arbitrary measurable space $(S, \mathcal{S})$. Our aim is to prove Theorem 1.12. The proof is quite different from that given in the countable case: it is based upon a detailed and somewhat technical analysis of the behavior of $Q(x_1, \cdots, x_n, E)$ as $n \to \infty$ for a typical sequence $(x_n) \in S^\infty$ and an exchangeable set E. The central part of the proof is Lemma 6.6, which establishes a certain property of $Q(x, E)$. This property is then shown to imply that $\mathcal{E}$ is trival: the arguments used may be useful for proving other kinds of zero-one law (see Remark 6.13).

We start with some preliminary lemmas. Let $(T, \mathcal{T})$ be another measurable space. Let A be a product measurable subset of $S \times T$. Let

$$_xA = \{y : (x,y) \in A\} \subset T, \qquad x \in S,$$

$$A_y = \{x : (x,y) \in A\} \subset S, \qquad y \in T.$$

Let λ and μ be probabilities on $(S, \mathcal{S})$ and $(T, \mathcal{T})$ respectively.

(6.1) LEMMA. *If* $\mu\{y : 0 < \lambda(A_y) < 1\} = 0$ *then*

$$\lambda\{x : \mu(_xA) = \lambda \times \mu(A)\} = 1.$$

The easy proof is left to the reader (verify that A coincides, up to some product null set, with $S \times \{y \in T : \lambda(A_y) = 1\}$).

(6.2) LEMMA. *Let* $0 < u \leqslant \frac{1}{2}$. *Then*

$$\lambda\{x : \mu({}_xA) < u\} \wedge \mu\{y : \lambda(A_y^c) < u\} < 2u.$$

PROOF. Put $P = \lambda \times \mu$, and denote the projections by X and Y. Let $f(x) = \mu({}_xA)$, $F = \{f(x) < u\}$, $g(y) = \lambda(A_y^c)$, $G = \{g(Y) < u\}$. We must prove

(6.3) $$P(F) \wedge P(G) < 2u.$$

But $uP(F) > P(FA) \geqslant P(AG) - P(F^cG) > (1-u)P(G) - P(F^c)P(G)$. This yields after rearrangement

$$(P(F) - 2u)(P(G) - 2u) + u(P(F) - 2u) + u(P(G) - 2u) < 0$$

and (6.3) follows.

Now let ϕ be a reference probability on $(S, \mathbb{S})$, fixed for the rest of the section.

(6.4) LEMMA. *Let A be an exchangeable subset of $S \times S$, μ a probability on $(S, \mathbb{S})$. Let $B \in \mathbb{S}$, $\varepsilon > 0$. Then*

$$\phi\{x : |\mu(A_x) - 1_B(x)| \geqslant \tfrac{1}{2}r(B, \varepsilon)\} > \varepsilon,$$

where $r(B, \varepsilon) = \inf\{\mu(B \setminus C) \wedge \mu(B^c \setminus C) : \phi(C) \leqslant \varepsilon\}$.

PROOF. By definition, A is exchangeable iff A is product measurable and $A_x = {}_xA$ for all $x \in S$. Thus taking $\lambda = \mu$ in Lemma 6.2 shows that for an exchangeable set A

(6.5) $$\mu\{x : \mu(A_x) < u\} \wedge \mu\{x : \mu(A_x^c) < u\} < 2u.$$

Now given B and ε with $r(B, \varepsilon) > 0$, let $u = \frac{1}{2}r(B, \varepsilon)$, $C = \{x : |\mu(A_x) - 1_B(x)| \geqslant u\}$. So $B^c \setminus C \subset \{x : \mu(A_x) < u\}$, $B \setminus C \subset \{x : \mu(A_x^c) < \mu\}$. If $\phi(C) \leqslant \varepsilon$ then $\mu(B \setminus C) \wedge \mu(B^c \setminus C) \geqslant 2u$ by definition of $r(B, \varepsilon)$, and this contradicts (6.5).

Now let (μ_n) be a given sequence of probabilities on $(S, \mathbb{S})$. Let Q be the product measure $\Pi_n \mu_n$ on $(S^\infty, \mathbb{S}^\infty)$, $\mathcal{E} \subset \mathbb{S}^\infty$ the exchangeable σ-field. For $A \in \mathbb{S}^\infty$, $x_1, \cdots, x_n \in S$ let $Q(x_1, \cdots, x_n, A)$ be defined as in (3.10), and note that (3.11) still applies. Let $r_n(B, \varepsilon)$ be as in (1.11), i.e., the value of $r(B, \varepsilon)$ in (6.4) for $\mu = \mu_n$. Here is the central result in the proof of Theorem 1.12.

(6.6) LEMMA. *Let $E \in \mathcal{E}$, $B \in \mathbb{S}$. Suppose $\varepsilon > 0$ is such that $\Sigma_n r_n(B, \varepsilon) = \infty$. Then*

$$\phi\{x : |Q(x, E) - 1_B(x)| \geqslant \tfrac{1}{2}\varepsilon\} \geqslant \tfrac{1}{2}\varepsilon.$$

PROOF. Since the definition of $Q(x, E)$ does not depend on μ_1, we assume $\mu_1 = \phi$. Observe first that

$$|Q(x, E) - 1_B(x)| = Q(x, D)$$

where D is the symmetric difference of $(X_1 \in B)$ and E, so it suffices to show that $Q(D) \geqslant \varepsilon$. Now for $A \in \mathbb{S}^\infty$ write $Q_{\mathcal{F}n}(A)$ for $Q(X_1, \cdots, X_n, A)$, which is a version of $Q(A|\mathcal{F}_n)$. Consider for each $n \geq 2$ the random variable

(6.7) $$H_n = Q_{\mathcal{F}_{n-1}}\big(Q_{\mathcal{F}_n}(D) \geq \tfrac{1}{2}\big) = h(X_1, \cdots, X_{n-1}),$$

where h can be specified as follows: for $\mathbf{y} = (y_2, \cdots, y_{n-1})$, $h(x, \mathbf{y}) = |q(x, \mathbf{y}) - 1_B(x)|$, where $q(x, \mathbf{y}) = \mu_n\{z : Q(x, \mathbf{y}, z, E) \geq \frac{1}{2}\}$. For each fixed $\mathbf{y}$, the symmetry of Q and Lemma 6.4 imply that for $a_n = \frac{1}{2}r_n(B, \varepsilon)$

$$\phi\{x : h(x, \mathbf{y}) \geq a_n\} \geq \varepsilon.$$

Because $\mu_1 = \phi$ this implies

(6.8) $$Q(H_n \geq a_n) \geq \varepsilon, \qquad n \geq 2, \qquad \text{where} \quad \Sigma_n a_n = \infty.$$

But an elementary argument shows that (6.8) and the positivity of the random variables H_n imply

$$Q(\Sigma_n H_n = \infty) \geq \varepsilon.$$

Now (6.7) and the conditional Borel-Cantelli lemma (Breiman (1968), Corollary 5.29) yield

$$Q\big(Q_{\mathcal{F}_n}(D) \geq \tfrac{1}{2} \text{ infinitely often}\big) \geq \varepsilon.$$

Hence $Q(D) \geq \varepsilon$ by martingale convergence.

(6.9) LEMMA. *Let $E \in \mathcal{E}$. If (μ_n) is ϕ-tame, then*

$$\mu_1\{x : Q(x, E) = Q(E)\} = 1.$$

PROOF. Let S_2^∞ be the product of all copies of S excluding the first, and let Q_2 be the probability $\Pi_2^\infty \mu_k$ on S_2^∞. Observe first that by martingale convergence, for $\phi \times Q_2$ almost all $(x, \mathbf{y}) \in S^\infty = S \times S_2^\infty$,

$$Q(x, \mathbf{y}_n, E) \to 1_E(x, \mathbf{y}) \qquad \text{as} \quad n \to \infty,$$

where $\mathbf{y} = (y_2, y_3, \cdots)$ and $\mathbf{y}_n = (y_2, \cdots, y_n)$. Thus for Q_2 almost all $\mathbf{y} \in S_2^\infty$

(6.10) $$\phi\{x : Q(x, \mathbf{y}_n, E) \to 1_E(x, \mathbf{y})\} = 1.$$

Now let $E_{\mathbf{y}} \subset S$ be the section of E at $\mathbf{y} \in S_2^\infty$, and define $F \subset S_2^\infty$ by

$$F = \{\mathbf{y} : 0 < \mu_1(E_{\mathbf{y}}) < 1\}.$$

By Lemma 6.1 it suffices to prove $Q_2(F) = 0$. Suppose, to obtain a contradiction, that $Q_2(F) > 0$. Choose and fix a $\mathbf{y} \in F$ such that (6.10) holds. Then

(6.11) $$\phi\{x : Q(x, \mathbf{y}_n, E) \to 1_B(x)\} = 1$$

where $B = \{x : (x, \mathbf{y}) \in E\}$ is such that $0 < \mu_1(B) < 1$. But because (μ_n) is ϕ-tame there exists $\varepsilon > 0$ such that $\Sigma r_n(B, \varepsilon) = \infty$. Thus Lemma 6.6 can be applied with $Q' = \mu_1 \times \Pi_{n+1}^\infty \mu_k$ substituted for Q and the section of E at $\mathbf{y}_n$ substituted for E to yield

(6.12) $$\phi\big\{x : |Q(x, \mathbf{y}_n, E) - 1_B(x)| \geq \tfrac{1}{2}\varepsilon\big\} \geq \tfrac{1}{2}\varepsilon.$$

There is now a contradiction between (6.11) and (6.12).

PROOF OF THEOREM 1.12. Suppose (μ_n) is ϕ-tame and E exchangeable. By Lemma 6.9, $Q(X_1, E) = Q(E)$ a.s. Now fix $\mathbf{y} = (x_1, \cdots, x_n)$ and apply Lemma 6.9 to $Q' = \Pi_{n+1}^{\infty} \mu_i$ and the section of E at $\mathbf{y}$ to obtain

$$Q(x_1, \cdots, x_n, X_{n+1}, E) = Q(x_1, \cdots, x_n, E) \qquad \text{a.s.}$$

Hence by induction $Q(X_1, \cdots, X_n, E) = Q(E)$ a.s., and martingale convergence shows $Q(E) = 0$ or 1.

(6.13) REMARK. The argument after Lemma 6.6 establishes the following result, which might be useful for proving zero-one laws in other circumstances.

PROPOSITION. *Let ϕ be a probability on $(S, \mathcal{S})$. Let Θ be a set of product measures on S^∞, and let $\mathcal{G}$ be a sub σ-field of $\mathcal{S}^\infty$. Suppose*

(i) $Q = \Pi \mu_n \in \Theta$ *implies* $Q' = \Pi_n \mu_{n+1} \in \Theta$;

(ii) $G \in \mathcal{G}$ *implies* $G' = \{(x_i) : (s, x_1, x_2, \cdots) \in G\} \in \mathcal{G}$, *each* $s \in S$;

(iii) *for each* $B \in \mathcal{S}$ *either*

(a) *for each* $Q = \Pi \mu_n \in \Theta$ *and each* $n \geq 1$, $\mu_n(B) = 0$ *or* 1;

or

(b) *there exists* $\varepsilon > 0$ *such that the conclusion of Lemma* 6.6 *holds for each* $Q \in \Theta, E \in \mathcal{G}$.

Then $Q(G) = 0$ *or* 1 *for each* $Q \in \Theta$, $G \in \mathcal{G}$.

Finally, we must prove Corollaries 1.13 and 1.14.

PROOF OF COROLLARY 1.13. We verify that (μ_n) is ϕ-tame for $\phi = \Sigma 2^{-n}\lambda_n$. Observe first that for any $n, j \geq 1$; $B, C \in \mathcal{S}$ we have

$$\mu_n(B \setminus C) \geq \lambda_j(B \setminus C) - \|\lambda_j - \mu_n\|$$
$$\geq \lambda_j(B) - 2^j\phi(C) - \|\lambda_j - \mu_n\|.$$

Using the same inequality for B^c,

$$r_n(B, \varepsilon) \geq \lambda_j(B) \wedge \lambda_j(B^c) - 2^j\varepsilon - \|\lambda_j - \mu_n\|.$$

Now let $B \notin \mathcal{S}_0$ be given. Then $0 < \lambda_j(B) < 1$ for some j. Since $\lambda_j \in \Lambda$,

$$\limsup_{n\to\infty} r_n(B, \varepsilon) \geq \lambda_j(B) \wedge \lambda_j(B^c) - 2^j\varepsilon.$$

For small enough $\varepsilon > 0$ this is positive, so $\Sigma r_n(B, \varepsilon) = \infty$.

PROOF OF COROLLARY 1.14. Again we verify that (μ_n) is ϕ-tame. If $B \notin \mathcal{S}_0$ then $0 < \phi(B) < 1$. Let $2\varepsilon = \phi(B) \wedge \phi(B^c)$. Then $\phi(B \setminus C) \wedge \phi(B^c \setminus C) \geq \varepsilon$ when $\phi(C) \leq \varepsilon$. But by the uniform absolute continuity, there exists $\delta > 0$ such that $\phi(A) \geq \varepsilon$ implies $\mu_n(A) \geq \delta$ for all n. So $r_n(B, \varepsilon) \geq \delta$ for all n.

7. Miscellaneous examples.

(7.1) EXAMPLE. Let $S = \{a, b, c, d\}$, and let

$$\mu_n(a) = \mu_n(b) = \tfrac{1}{2}, \qquad n \text{ even};$$
$$\mu_n(c) = \mu_n(d) = \tfrac{1}{2}, \qquad n \geqslant 3, \qquad n \text{ odd};$$
$$\mu_1(c) = \mu_1(a) = \tfrac{1}{2}.$$

Then (μ_n) does not satisfy Condition (a) although the requirements of that condition are satisfied by each singleton.

(7.2) EXAMPLE. Let $S = \{0, 1, 2\}$. Let $\mathcal{E}$ be the exchangeable σ-field on sequence space, and for $B \subset S$ let $\mathcal{E}_B$ be the exchangeable σ-field for $(1_B(X_n))$. Let $E = \{(x_i) : (x_{2i-1}, x_{2i}) = (0, 1)$ for infinitely many $i\}$. Clearly $E \in \mathcal{E}$. We shall describe a distribution for (X_i) under which each $\mathcal{E}_B$ is trivial, but $P((X_i) \in E) = \frac{1}{2}$. Hence $\mathcal{E}$ is strictly larger than $\bigvee_B \mathcal{E}_B$.

Let (Y_i) be an independent sequence distributed uniformly on S. Let α be independent of (Y_i), with $P(\alpha = 0) = P(\alpha = 1) = \frac{1}{2}$. Define

$$\begin{aligned} X_i &= Y_i + \alpha \quad \text{modulo } 3, \qquad \text{if } i \text{ is even and } \quad Y_i = Y_{i-1} + 1 \quad \text{modulo } 3; \\ &= Y_i \qquad \text{otherwise.} \end{aligned}$$

Then $\{(X_i) \in E\} = \{\alpha = 0, (Y_i) \in E\}$, and so $P((X_i) \in E) = \frac{1}{2}P((Y_i) \in E) = \frac{1}{2}$. But it may be verified that, for each $B \subset S$, the sequences $(1_B(X_i))$ and $(1_B(Y_i))$ have the same distribution. So $\mathcal{E}_B$ is trivial by the Hewitt-Savage result.

(7.3) EXAMPLE. Let S be the set of binary rationals in the interval $[0, 1]$. We shall describe (μ_n) so that Condition (a) is satisfied but $\mathcal{E}$ is not trivial. Given $n \geqslant 1$, write $n = 2^i + m$ for $i \geqslant 0$, $0 \leqslant m < 2^i$ and define

$$\mu_n(m2^{-i}) = \mu_n((m + 1)2^{-i}) = \tfrac{1}{2}.$$

By applying the argument of Theorem 1.1 to $V_n = X_n - EX_n$, observing that $EV_n^2 = 2^{-2i-2}$ so that $\Sigma EV_n^2 < \infty$, we see that $\mathcal{E}$ is not trivial. But consider a proper subset B of S. Then there exists $j \geqslant 1$ such that the points $\{m2^{-j}; 0 \leqslant m < 2^j\}$ are neither all in B nor all in B^c. So for each $i \geqslant j$ there exists $0 \leqslant m < 2^i - 1$ such that exactly one of the points $(m2^{-i}, (m + 1)2^{-i})$ is in B. So $\mu_{m+2^i}(B) = \frac{1}{2}$. Hence $\Sigma\, \mu_n(B) \wedge \mu_n(B^c) = \infty$, and Condition (a) is satisfied.

(7.4) EXAMPLE. Let $S = \{0, 1, 2, \cdots\}$. Define

$$\mu_n(0) = \mu_n(n) = \tfrac{1}{2}, \qquad n \geqslant 1.$$

Recall that $\mathcal{S}_+ = \{B \subset S : \Sigma\, \mu_n(B) \wedge \mu_n(B^c) < \infty\}$. In this case, $\{1, \cdots, n\} \in \mathcal{S}_+$ for each n, but $\{1, 2, \cdots\} \notin \mathcal{S}_+$. So $\mathcal{S}_+$ is not a σ-field.

(7.5) EXAMPLE. Let S be a countably infinite set, and $K \geqslant 2$. Let (μ_n) be an enumeration of all the probability measures uniform on K-point subsets of S. If

$K \geqslant 3$ then each pair of points is linked, so Condition (b) holds and $\mathfrak{S}$ is trivial by Theorem 1.8. Suppose now that $K = 2$. Then each distinct pair of points is weakly linked but not linked. So Condition (b) is not satisfied. Nevertheless we shall show that $\mathfrak{S}$ is trivial. By Proposition 3.13 it suffices to show that each distinct pair of points can be coupled.

Fix $y \neq z \in S$, $m \geqslant 1$. Write $n(0) = m$. Define $(x_i, j(i), k(i), n(i))$ inductively as follows. Choose $x_i \in S$ such that $\mu_j(x_i) = 0$ for each $j \leqslant n(i-1)$. Then we can choose $j(i), k(i) > n(i-1)$ such that

$$\mu_{j(i)}(x_i) = \mu_{j(i)}(y) = \tfrac{1}{2}.$$
$$\mu_{k(i)}(x_i) = \mu_{k(i)}(z) = \tfrac{1}{2}.$$

Let $n(i) = j(i) \vee k(i)$. We now describe a (m, y, z) coupling. Let $(Y_n, n \geqslant 2)$ have distribution $\prod_2^\infty \mu_n$. Let $Z_n = Y_n$ if $n \notin \cup_i \{j(i), k(i)\}$. Let

$$\begin{aligned}(Z_{j(i)}, Z_{k(i)}) &= (y, x_i) \quad \text{if} \quad (Y_{j(i)}, Y_{k(i)}) = (x_i, z)\\ &= (x_i, z) \quad \text{if} \quad (Y_{j(i)}, Y_{k(i)}) = (y, x_i)\\ &= (Y_{j(i)}, Y_{k(i)}) \qquad \text{otherwise.}\end{aligned}$$

The $(Z_n, n \geqslant 2)$ also has distribution $\prod_2^\infty \mu_n$. The remainder of the proof resembles that of Proposition 3.14. By construction, $M_{n(i)}(\mathbf{Y}, s) = M_{n(i)}(\mathbf{Z}, s)$ for $s \notin \{y, z\}$. So let

$$D_i = M_{n(i)}(\mathbf{Y}, y) - M_{n(i)}(\mathbf{Z}, y)$$
$$T = \inf\{n(i) : D_i = 0\}.$$

Then $M_T(\mathbf{Y}) = M_T(\mathbf{Z})$ on $\{T < \infty\}$. Defining $\mathbf{Z}^*$ as in (3.19) the required coupling is given by $(\mathbf{Y}, \mathbf{Z}^*)$ provided that $T < \infty$ a.s. To prove $T < \infty$ a.s., observe that (D_i) is a random walk with jump probabilities

$$P(D_i = D_{i-1} \pm 1) = \tfrac{1}{4}$$
$$P(D_i = D_{i-1}) = \tfrac{1}{2}.$$

REFERENCES

[1] Blum, J. R. and Pathak, P. K. (1972). A note on the zero-one law. *Ann. Math. Statist.* **43** 1008–1009.

[2] Brieman, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.

[3] Freedman, D. (1971). *Markov Chains*. Holden-Day, San Francisco.

[4] Georgii, H. O. (1976). On canonical Gibbs states, symmetric and tail events. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **33** 331–341.

[5] Griffeath, D. (1975). A maximal coupling of Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **31** 95–106.

[6] Hewitt, E. and Savage, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501.

[7] Horn, S. and Schach, S. (1970). An extension of the Hewitt-Savage zero-one law. *Ann. Math. Statist.* **41** 2130–2131.

[8] LÉVY, P. (1937). *Théorie de l'Addition des Variables Aléatoires*. Bautier-Villars, Paris.

[9] MINEKA, J. (1973). A criterion for tail events for sums of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **25** 163–170.

[10] OREY, S. (1966). Tail events for sums of independent random variables. *J. Math. Mech.* **15** 937–951.

[11] PITMAN, J. W. (1976). On coupling of Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **35** 315–322.

[12] PITMAN, J. W. (1978). An extension of de Finetti's Theorem. Summary to appear in *Advnaces in Applied Probability, Proceedings of the Seventh Conference on Stochastic Processes and their Applications*.

[13] SENDLER, W. (1975). A note on the proof of the zero-one law of Blum and Pathak. *Ann. Probability* **3** 1055–1058.

[14] SIMONS, G. (1978). Some extensions of the Hewitt-Savage zero-one law. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **42** 167–173.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
16 MILL LANE
CAMBRIDGE CB2 ISB, ENGLAND