Relations on Probability Spaces and Arrays of Random Variables

Douglas N. Hoover

Institute for Advanced Study
School of Mathematics
Princeton, NJ 08540

Abstract

We discuss a generalized model for probability suitable for representing relations between members of populations. The study of such relations via a certain sampling procedure is equivalent to studying multiply indexed arrays of random variables $\{X_{i_1,\ldots,i_n}\}_{i_1,\ldots,i_n \in \mathbb{N}}$ whose distributions are invariant under permutations taking $i_1,\ldots,i_n$ to $\pi(i_1),\ldots,\pi(i_n)$, π a permutation of $\mathbb{N}$. Relations are also closely related to arrays satisfying similar partial exchangeability conditions. By a measure algebraic analysis of relations we show that such arrays as we have mentioned may be represented as functions of i i.d. or equally simple arrays. This implies a generalization of a recent theorem of D. J. Aldous.

1. Introduction

A random variable may be thought of as a measurement on some population. The moving idea of this paper is to consider and analyse probabilistically relations among individual members of populations. We do this using the following framework, invented by H. J. Keisler ([8]). (Our whole point of view on relations is a development of his ideas).

1.1. Definition: (a) A graded probability space (g.p.s.) is a structure $(\Omega, \mathcal{F}_n, P_n)_{n\in\mathbb{N}}$ where for each n, $(\Omega^n, \mathcal{F}_n, P_n)$ is a probability space, and this sequence of probability spaces satisfies the Fubini Property (namely, the conclusions of Fubini's Theorem), i.e., for each $m, n \in \mathbb{N}$,

(i) If π is a permutation of $\{1, \ldots, n\}$, $F \in \mathcal{F}_n$, then

$$F^\pi = \{(\omega_{\pi(1)}, \ldots, \omega_{\pi(n)}) : (\omega_1, \ldots, \omega_n) \in F\} \in \mathcal{F}_n$$

and

$$P_n(F) = P_n(F^\pi) .$$

(ii) $\mathcal{F}_n \times \mathcal{F}_m \subseteq \mathcal{F}_{n+m}$

(iii) If $F \in \mathcal{F}_{n+m}$, and $\vec{\omega} \in \Omega^m$, then

$$F(\cdot, \vec{\omega}) = \{\vec{\omega}' \in \Omega^n : (\vec{\omega}', \vec{\omega}) \in F\} \in \mathcal{F}_n ,$$

the function $g(\vec{\omega}) = P_n(F(\cdot, \vec{\omega}))$ is $\mathcal{F}_m$-measureable, and

$$\int g(\vec{\omega}) dP_n(\vec{\omega}) = P_{n+m}(F) .$$

(b) An (n-dimensional) relation on Ω, or a relation on Ω^n, is an $\mathcal{F}_n$-measureable random variable on Ω^n (with values in some complete separable metric space). ■

The obvious example of a g.p.s. is $(\Omega, \mathcal{F}^n, P^n)_{n\in\mathbb{N}}$ where $(\Omega, \mathcal{F}, P)$ is a probability space, and $\mathcal{F}^n, P^n$ are the usual product σ-algebras and probabilities. The idea of allowing $\mathcal{F}_{n+m}$ to be bigger than $\mathcal{F}_n \times \mathcal{F}_m$ is to leave room for more general relations. That there is some advantage in this should become clear in the course of the paper.

With a single random variable, there is, from a probabilistic point of view, really nothing of interest but the distribution. A relation is a more complex object. One would like to know its symmetries (how it varies under permutation of the coordinates of Ω^n), how its sections $X(\cdot, \vec{\omega})$, $\vec{\omega} \in \Omega^m$, $m < n$, are related, etc. The simplest way to find out about these things to look at the distribution of various arrays of random variables obtained by a sampling procedure analogous to the method of producing a sequence of i. i. d. random variables from a single random variable X by letting

$$X_n(\omega) = X(\omega(n)) \qquad \omega \in \Omega^{\mathbb{N}} .$$

(There are more complicated ways expressing probabilistic facts about a relation, but they come to the same thing; see [7]).

1. 2. Definition: (a) For $n \in \mathbb{N}$, we will also let n denote the set $\{1, \ldots, n\}$. In the same spirit, we will let $n \times m$ denote the Cartesian product of the sets denoted by the two numbers (and similarly for $\mathbb{N} \times n$), and $n!$ to denote the permutations of n, $\mathbb{N}!$ those of $\mathbb{N}$. $(\mathbb{N}\times n)!$ denotes the restricted permutations of $\mathbb{N} \times n$, namely those of the form

$$\pi(m, i) = (\pi_i(m), i) \qquad m \in \mathbb{N},\ i \leq n$$

for $\pi_i \in \mathbb{N}!$

(b) $\mathcal{F}_{\mathbb{N}}$ is the smallest σ-algebra on $\Omega^{\mathbb{N}}$ such that $\tilde{G} = \{\omega \in \Omega^{\mathbb{N}} : (\omega_{i_1}, \ldots, \omega_{i_n}) \in G\}$ is $\mathcal{F}_{\mathbb{N}}$ measurable and $P_{\mathbb{N}}$ is the unique probability on

$\mathcal{F}_{\mathbb{N}}$ such that $P_{\mathbb{N}}(\tilde{G}) = P_n(G)$ for each $n, i_1, \ldots, i_n \in \mathbb{N}, G \in \mathcal{F}_n$. Define $\mathcal{F}_{\mathbb{N}\times n}$ and $P_{\mathbb{N}\times n}$ on $\Omega^{\mathbb{N}\times n}$ similarly.

(c) we will use the letters σ, τ to denote finite sequences of natural numbers, i.e., members of $\mathbb{N}^n$, for some n. If $\sigma \in \mathbb{N}^n$, $\pi \in \mathbb{N}!$, σ^{π} is the sequence $(\pi(\sigma(1)), \ldots, \pi(\sigma(n)))$. If $\pi \in (\mathbb{N}\times n)!$,

$$\sigma^{\pi} = (\pi_1(\sigma(1)), \ldots, \pi_n(\sigma(n))) .$$

we will use s, t, u, v to denote sets of natural numbers and define s^{π}, $\pi \in \mathbb{N}!$, similarly.

(d) Let X be a relation on Ω^n, $(\Omega, \mathcal{F}_m, P_m)_{m\in\mathbb{N}}$ a g.p.s.

(i) $\{X_{\sigma}^{DRE}\}_{\sigma\in\mathbb{N}^n}$ is the array of random variables on $\Omega^{\mathbb{N}}$ which X induces by simple sampling; i.e.,

$$X_{\sigma}(\omega) = X(\omega(\sigma(1)), \ldots, \omega(\sigma(n))), \quad \sigma \in \mathbb{N}^n .$$

(ii) $\{X_{\sigma}^{ED}\}_{\sigma\in\mathbb{N}^n}$ is the array of random variables on $\Omega^{\mathbb{N}\times n}$ induced from X by ramified sampling, i.e.,

$$X_{\sigma}(\omega) = X(\omega(\sigma(1), 1), \ldots, \omega(\sigma(n), n)), \qquad \sigma \in \mathbb{N}^n . \qquad \blacksquare$$

$\{X_{\sigma}^{ED}\}$ (we will explain the superscripts DRE and ED presently) can tell us rather less about X than $\{X_{\sigma}^{DRE}\}$ can; for instance one cannot derive from it the joint distributions of $(X^{\pi} : \pi \in n!)$ (where X^{π} is defined analogously to F^{π} in 1.1(a)(i)), though one can from $\{X_{\sigma}^{DRE}\}$. The ramified sampling would be particularly appropriate when the symmetries of X were not of interest, for instance if the different coordinates of Ω^n are meant to represent different kinds of things not to be compared with one another. But in any case it is useful because it isolates the structure of X which is relevant to dimension (see §5).

Now $\{X_\sigma^{DRE}\}$ has the symmetry property

$$(1) \qquad \{X_\sigma^{DRE}\} \sim \{X_{\sigma\pi}^{DRE}\} \qquad \pi \in \mathbb{N}!$$

(where $\sim$ indicates that the two arrays have the same distribution), whereas $\{X_\sigma^{ED}\}$ has the stronger property

$$(1)' \qquad \{X_\sigma^{ED}\} \sim \{X_{\sigma^\pi}^{ED}\} \qquad \pi \in (\mathbb{N}\times n)!$$

(this is stronger since it allows different permutations of $\mathbb{N}$ to be applied to each place in a sequence σ). Both have the further property that

$$(2) \qquad \{X_\sigma\}_{\sigma\in m^n} \text{ is independent of } \{X_\sigma\}_{\sigma\in(\mathbb{N}-m)^n}\,.$$

(All this follows directly from the Fubini Property.) Now these properties have been considered before: A distribution with property (1) is essentially what was called a symmetric measure model by Gaifman [4]; (1) plus (2) is the notion of symmetric product measure of Krauss [9]. Arrays with the property (1)' have been called row column exchangeable (RCE for short, see Aldous [1]), those with (1)' plus (2) exchangeably disassociated (ED--see Silverman [14]). In conformity with the latter terminology we will call property (1) relationally exchangeable (RE) and (1) plus (2) dissociated relationally exchangeable (DRE).

Consider also a third type of array: $\{Y_s : s \subseteq \mathbb{N}, |s| = n\}$ is said to be weakly exchangeable (WE; see Eagleson and Weber [3]) if $\{Y_s\} \sim \{Y_{s\pi}\}$, $\pi \in \mathbb{N}!$. Such arrays may also be obtained (essentially) by sampling relations: for suppose $X : \Omega^n \to \mathbb{R}$ has $X^\pi = X$, $\pi \in n!$. Then when we form $\{X^{DRE}\}$, if σ is one-to-one, X_σ^{DRE} does not depend on the order of $(\sigma(1), \ldots, \sigma(n))$ but only on the set $\{\sigma(1), \ldots, \sigma(n)\}$. Thus, if we define $\{X_s^{DWE}\}$ by $X_s^{DWE} = X_\sigma^{DRE}$, where $s = \{\sigma(1), \ldots, \sigma(n)\}$, this array is weakly exchangeable, and also dissociated in a sense similar to (2) (hence DWE).

The study of relations as we propose it, reduces then to the study of arrays with the exchangeability and dissociation properties we have outlined. In fact, there is an equivalence, because every DRE, ED, or DWE array has the same distribution as our array induced by a relation on a g.p.s. via the corresponding sampling procedure (Theorem 4.1). From the point of view of studying these kinds of arrays, the formulation as a relation on a g.p.s. provides a sort of concreteness and an appropriate framework that makes many things obvious which would not otherwise be so. It shows the single concept on which ED, DRE and DWE vary.

As we said in §4 we show that every ED, DRE, or DWE array is induced by sampling. This requires a special technique for constructing probability spaces, the method of ultraproducts. In our context this construction amounts to a method of forming limits of probability spaces which coordinates properly with weak convergence of random variables. This is contained in §3.

The main aim of this paper is to prove a theorem about the measure algebraic structure of relations which, transferred to arrays, implies a generalization to n indices (and to DRE arrays) of the recent theorem of Aldous ([1]) which states that if $\{X_{ij}\}_{i,\, j \in \mathbb{N}}$ is ED, then it may be functionally represented as $\{X_{ij}\} \sim \{f(\xi_i, \eta_j, U_{ij})\}$ where f is Borel measureable and the family $\{\xi_i, \eta_j, U_{ij} : i, j \in \mathbb{N}\}$ is i.i.d. §5 contains results about how relations interact with relations of lower dimension (generalizing Lemma 3.6 of [1]), and §6 some fairly simple, measure algebraic results, which are needed as technical tools. The theorem itself is proved in §7. There we also characterize when two functional representations induce the same distribution; this answers a question of Aldous.

Versions of all our results follow for RCE, RE, and WE arrays (i.e., without dissociation). This requires only a simple device: any RCE (or DRE or WE) array $\{X_\sigma\}_{\sigma \in \mathbb{N}^n}$ can be considered to be $\{Y_{(\sigma, 1)}\}$ for an array $\{Y_\tau\}_{\tau \in \mathbb{N}^{n+1}}$ which is ED (DRE, DWE).

An earlier, more primitive version of most of the results in this paper was announced in [6]. This version, especially §5, has been improved by ideas from Aldous' paper.

2. Further Preliminaries.

Notation, conventions, and some results we will use.

2.1. Notation and conventions:

(a)(i) We will be using many different probabilities in this paper, and at least some of the time it will be useful to distinguish their expectations. Therefore if P is a probability, and X a real random variable, $P(X)$ will denote the expectation of X with respect to P, and if $\mathcal{G}$ is a σ-field, $P(X|\mathcal{G})$ is the conditional expectation with respect to $\mathcal{G}$ (in the P-measure).

(ii) If X is an M-valued random variable, $\mathcal{G}$ a σ-field, P a probability, then $\mathcal{D}(X|\mathcal{G})$ is the conditional distribution of X with respect to $\mathcal{G}$, i.e., a $\mathcal{G}$-measureable random measure such that

$$\mathcal{D}(X|\mathcal{G})(\varphi) = P(\varphi(X)|\mathcal{G}) \quad \text{a.s.}$$

for any bounded continuous $\varphi : M \to \mathbb{R}$. (Of course, this notation is inconsistent with (i).)

(iii) If X and Y are random variables, $X \sim Y$ means they have the same distribution.

(b) Let $(\Omega, \mathcal{F}, P)$ be a probability space

(i) If X is an $\mathcal{F}$-measureable random variable, $\mathcal{F}(X)$ denotes the σ-field generated by X

(ii) If $\mathcal{G}$ is a sub-σ-field of $\mathcal{F}$, and $F_i \in \mathcal{F}$ for each $i \in I$, then $\mathcal{G}(F_i : i \in I)$ is the smallest σ-field containing $\mathcal{G}$ and $\{F_i : i \in I\}$.

(iii) U^c denotes the complement of the subset U of Ω

(iv) If $\mathcal{F}$ is a field, $\sigma(\mathcal{F})$ denotes the σ-field generated by $\mathcal{F}$.

(c) Conventions: we use numbers to denote sets:

(i) $n = \{1, \ldots, n\}$

$n \times m = \{(i, j) : i \leq n, \ j \leq m\}$, the cartesian product of n and m (likewise $\mathbb{N} \times n$).

n^m is the m-fold Cartesian product of n with itself (the sequences from n of length m).

(ii) $P(n, m)$ denotes the set of 1-1 sequences from n of length m.

$C(n, m)$ denotes the subsets of n of cardinality m. $P(\mathbb{N}, m)$, $P(\mathbb{N}\times n, m)$ etc. are defined similarly.

(iii) $n!$ denotes the set of permutations of n, $\mathbb{N}!$ those of $\mathbb{N}$, $(\mathbb{N}\times n)!$ the restricted permutations of $\mathbb{N} \times n$ (i.e., essentially $(\mathbb{N}!)^n$). A permutation π is said to be finite if it fixes all but finitely many elements.

(iv) The letters s, t, u will denote subsets of $\mathbb{N}$ or $\mathbb{N} \times n$. The notations s^m, $P(s, m)$, $s!$, etc., are extended to these in the obvious way.

(v) Let $(\Omega, \mathcal{F}_n, P_n)_{n\in\mathbb{N}}$ be a g.p.s. For $s \subseteq \mathbb{N}$ or $\mathbb{N} \times n$, Ω^s is the space of sequences from Ω indexed by s, $\mathcal{F}_s$ is the smallest σ-algebra on Ω^s such that for each $i_1, \ldots, i_n \in s$, $n \leq |s|$,

$$\tilde{G} = \{\omega \in \Omega^s : (\omega(i_1), \ldots, \omega(i_n)) \in G\} \in \mathcal{F}_s$$

for every $G \in \mathcal{F}_n$. P_s the unique probability such that $P_s(\tilde{G}) = P_n(G)$ for each such G. P_s is countably additive by A. Ionescu-Tulcea's Theorem.

(vi) If $\omega \in \Omega^S$, $\pi \in s!$, then ω^π is given by

$$\omega^\pi(i) = \omega(\pi(i)) \qquad i \in I .$$

$$F^\pi = \{\omega^\pi : \omega \in F\} \qquad F \in \mathcal{F}_S .$$

But for $\sigma \in \mathbb{N}^S$, $\pi \in \mathbb{N}!$,

$$\sigma^\pi(i) = \pi(\sigma(i)) \quad (\text{or } \pi_i(\sigma(i)), \text{ if } \pi \in (\mathbb{N}\times s)!), \ i \in s.$$

(vii) We will assume that all relations on Ω take values in [0, 1]. There is no loss of generality in this, since we are dealing only with distributions, and every complete separable metric space is Borel isomorphic to [0, 1]. ∎

DeFinetti's Theorem says that the distribution of any exchangeable sequence of random variables (i.e., X_n $n \in \mathbb{N}$ such that $\{X_n\} \sim \{X_{\pi(n)}\}$, for every $\pi \in \mathbb{N}!$) is a mixture of i.i.d. distributions. For arrays the properties RCE, RE, and WE are analogs of exchangeability which take into account the array structure, and ED, DRE, DWE are the respective analogues of i.i.d. Thus we have the following analogue of de Finetti's Theorem.

2.2. Theorem: Let $\{X_\sigma\}_{\sigma\in\mathbb{N}^n}$ be an RCE array, with values in a complete separable metric space M

(i) $P(\{X_\sigma\} \in B) = \int Q(B)d\mu(Q)$ for each Borel set $B \subseteq M^{\mathbb{N}^n}$, where μ is a probability on (the weak*-Borel sets of) the ED distributions Q.

(ii) This μ is unique.

(iii) If $\mathcal{J} = \bigcap_m \mathcal{F}\{X_\sigma : \sigma \in (\mathbb{N}-m)^n\}$, then $\mathcal{D}(\{X_\sigma\}|\mathcal{J})$ is ED almost surely, and its distribution is the same as μ.

The obvious analogues hold for RE/DRE and WE/DWE. ∎

A proof of this theorem can be obtained by an almost mechanical adaptation of any of the usual proofs of de Finetti's Theorem (as may be found in Hewitt-Savage [5] or Loeve [11]). A proof of what is essentially the RE/DRE case of the theorem is found in Krauss [9], but we think changes needed are small enough for the proof of 2.2 to be omitted.

The Hewitt-Savage zero-one law extends immediately to ED, DRE, and DWE arrays.

2.3. Theorem. Suppose X is ED (resp. DRE) for $m \in \mathbb{N}$, let ξ_m be the σ-algebra of sets of the form $\{\omega : \{X_\sigma\} \in B\}$, where $B \subseteq M^{\mathbb{R}^n}$ is Borel and

$$\{\omega:\{X_\sigma\} \in B\} = \{\omega:\{X_{\sigma\pi}\} \in B\}$$

for every $\pi \in \mathbb{N} \times n!$ (resp. $\mathbb{N}!$) which fixes every element of $(\mathbb{N}-m) \times n$ (resp. $\mathbb{N}-m$). Then for every $E \in \xi = \bigcap_n \xi_n$, $P(E)$ is zero or one. ∎

This follows by the proof of the usual Hewitt-Savage zero-one law, with no essential change.

3. Ultraproducts of Graded Probability Spaces

The results in this section are just a treatment of certain facts we need from logic and nonstandard analysis (namely, parts of the Los Ultraproduct Theorem, Loeb's theorem on standardizing nonstandard measures, and an observation of Keisler on Fubini's Theorem for Loeb measures of nonstandard product measures), which we mean to make accessible to the probabilistic reader who is not familiar with either logic or nonstandard analysis. Fuller treatments of these three topics may be found in Chang and Keisler [2], Loeb [10], and Keisler [8]. The Ultraproduct Theorem particularly is more elegant given a prior development of first order logic.

3.1. Definition: Let I be a set. An ultrafilter on I is a set D of subsets of I, such that

(i) For $U_1, \ldots, U_n \in D$, $U_1 \cap \ldots \cap U_n \neq \phi$.

(ii) For each $U \subseteq I$, either $U \in D$ or $U^c \in D$. ∎

3.2. Proposition: (a) If D is an ultrafilter then

(i) $U \in D$, $V \supseteq U$ implies $V \in D$.

(ii) If $U, V \in D$, then $U \cap V \in D$.

(b) Any set of subsets of I satisfying 3.1(i) is contained in an ultrafilter on I.

Proof: (a) is easy; (b) is easy using Zorn's Lemma. ∎

3.3. Definition: Let I be an index set, D an ultrafilter on I.

(i) If $r \in [0,1]^I$ then

$$r^D = (r(i))^D = \sup\{r' \in [0,1] : \{i : r(i) > r'\} \in D\}.$$

(ii) If G is a set, then $G^D = \prod_D G$, the ultrapower of G, is the product G^I modulo the equivalence relation

$$f \approx_D g \text{ iff } \{i : f(i) = g(i)\} \in D \qquad (f, g \in G^I).$$

$\approx_D$ is obviously symmetric and reflexive; it is transitive, since if $f \approx_D g$ and $g \approx_D h$, then

$$\{i : f(i) = g(i)\} \supseteq \{i : f(i) = g(i)\} \cap \{i : g(i) = h(i)\},$$

so $\{i : f(i) = h(i)\} \in D$ by (a) of Prop. 3.2. Write f^D for the equivalence class of f modulo $\approx_D$.

(iii) If $\lambda(i) : G \to [0,1]$ for each $i \in I$, then $\lambda^D : G^D \to [0,1]$ is defined by

$$\lambda^D(f^D) = (\lambda(i)(f(i)))^D .$$

That this definition is independent of the representative of f^D used on the right again follows easily by Prop. 3.2(a).

(iv) If $F(i) \subseteq G$ for each $i \in I$, then $(F(i))^D$ is the subset F of G^D whose characteristic function χ_F is $(\chi_{F(i)})^D$, the ultraproduct of the characteristic functions of the F(i)'s. Thus

$$f^D \in (F(i))^D \text{ iff } \{i : f(i) \in F(i)\} \in D .$$

Since $F^D = G^D$ iff $\{i : F(i) = G(i)\} \in D$, ultraproducts of sets of sets may be defined similarly.

(v) Given any sets $F(i)$, $i \in I$, we can define their ultraproduct $(F(i))^D = F^D$ by considering $F(i)$ as subsets of $\bigcup_{i \in I} F(i)$; likewise we can define ultraproducts of functions, sets, and sets of sets on $F(i)$, $i \in I$. ∎

3.4. Lemma: For $r \in [0,1]^I$

$$r^D = \inf\{q \in [0,1] : \{i : q > r(i)\} \in D\}$$

Proof: Let $q > r^D$. Then $\{i : r(i) > q - \varepsilon\} \notin D$, for some $\varepsilon > 0$. Hence $\{i : r(i) \geq q\} \notin D$, and so $\{i : q > r(i)\} \in D$ (the last two inferences both by Prop. 3.2(a)). ∎

3.5. Proposition: If $\varphi \in C([0,1]^n)$, $r_i, \ldots, r_n \in [0,1]^I$, then

$$\varphi(r_1^D, \ldots, r_n^D) = (\varphi(r_1(i), \ldots, r_n(i)))^D .$$

Proof: Fix $\varepsilon > 0$, and choose $\delta > 0$ such that if $\max_{j \leq n} |q_j - r_j^D| < \delta$ $q_1, \ldots, q_n \in [0,1]$, then $|\varphi(\vec{q}) - \varphi(\vec{r}^D)| < \varepsilon$. By the definition of r^D, Lemma 3.4, and Prop. 3.2(a), $\{i : |r_j(i) - r_j^D| < \delta, 1 \leq j \leq n\} \in D$. Hence

$$\{i : |\varphi(r_1(i), \ldots, r_n(i)) - \varphi(r_1^D, \ldots, r_n^D)| < \varepsilon\} \in D,$$

and the proposition follows. ∎

Before proceeding to g.p.s.'s, we will discuss ultraproducts of probability spaces. The ultraproduct of a g.p.s. $(\Omega, \mathcal{F}_n, P_n)_{n \in \mathbb{N}}$ is obtained by just putting together the ultraproducts of its layers $(\Omega^n, \mathcal{F}_n, P_n)$, $n \in \mathbb{N}$.

3.6. Definition: (a) An ultrafilter D on I is said to be countably incomplete if there are $U_n \in D$, $n \in \mathbb{N}$, such that $\bigcap_n U_n = \phi$.

(b) Let $(\Omega(i), \mathcal{F}(i), P(i))$ $i \in I$, be probability spaces, D a countably incomplete ultrafilter on I. The ultraproduct of $(\Omega(i), \mathcal{F}(i), P(i))$ with respect to D is the probability space $(\Omega, \mathcal{F}, P)$, where $\Omega = (\Omega(i))^D$, $\mathcal{F} = \sigma((\mathcal{F}(i))^D)$, and P is the unique extension of P^D to $\mathcal{F}$. ∎

We can see that countably incomplete ultrafilters exist on any infinite index set I, for if we write $I = \bigcup_n I_n$ as the countable union of nonempty disjoint sets, then the set $\{U_m : m \in \mathbb{N}\}$, $U_m = \bigcup_{n \geq m} I_n$, can be extended to an ultrafilter, which must be countably incomplete since $\bigcap_m U_m = \phi$.

The following proposition shows that part (b) of this definition actually makes sense:

3.7. Proposition: (a) $\mathcal{F}^D$ is a field of sets.

(b) P^D is a finitely additive probability measure on $\mathcal{F}^D$.

(c) (Loeb's Theorem [10]). P^D is countably additive on $\mathcal{F}^D$, hence has a unique extension to $\sigma(\mathcal{F}^D)$.

Proof: (a) Let $F = (F(i))^D \in \mathcal{F}^D$. By Prop. 3.5,

$$\chi_{F^c} = 1 - \chi_F = (1-\chi_{F(i)})^D = (\chi_{F(i)^c})^D ,$$

hence $F^c = (F(i)^c)^D \in \mathcal{F}^D$. Similarly $F \cap G \in \mathcal{F}^D$ for $F, G \in \mathcal{F}^D$.

(b) Prop. 3.5 also implies that $F, G \in \mathcal{F}^D$ are disjoint iff $\{i : F(i) \cap G(i) = \phi\} \in D$, and that in this case

$$\begin{aligned} P^D(F) + P^D(G) &= (P(i)(F(i)) + P(i)(G(i)))^D \\ &= (P(i)(F(i) \cup G(i)))^D \\ &= P^D(F \cup G). \end{aligned}$$

(c) If we show P^D is countably addive on $\mathcal{F}^D$, then P^D will have a unique extension to $\sigma(\mathcal{F}^D)$ by the Caratheodory-Hopf extension Theorem. It suffices to show that if $F_n \in \mathcal{F}^D$, $F_n \supseteq F_{n+1}$, $n \in \mathbb{N}$, and $\bigcap_n F_n = \phi$, then $P^D(F_n) \to 0$. We will show that in fact $\bigcap_n F_n = \phi$ can happen only if $F_m = \phi$ for some $m \in \mathbb{N}$, which will suffice to prove the result. For suppose $F_n \neq \phi$, $n \in \mathbb{N}$. Choose $\omega_n \in \prod_{i \in I} \Omega(i)$ such that

$$S_n = \{i : \omega_n(i) \in F_n(i)\} \in D .$$

Let $U_n \in D$, $n \in \mathbb{N}$, $\bigcap_n U_n = \phi$, and let

$$T_n = \bigcap_{k \leq n} (S_k \cap U_k) .$$

Define $\omega(i)$ by

$$\omega(i) = \omega_n(i) \qquad \text{for } i \in T_n - T_{n+1}$$
$$= \omega_1(i) \qquad i \notin T_1 .$$

Since T_n are decreasing and $\bigcap_n T_n = \phi$, $\omega(i)$ is well-defined for each $i \in I$. Now

$$\{i : \omega(i) \in F_n(i)\} \supseteq T_n \in D$$

for each $n \in \mathbb{N}$, so $\omega^D \in \bigcap_n F_n^D$. ∎

Now we connect ultraproducts with weak convergence.

3.8. Proposition: (a) Suppose $X(i)$, $i \in I$, are $\mathcal{F}(i)$-measureable random variables on $\Omega(i)$. Then X^D is $\mathcal{F}$-measureable and $P(X^D) = (P(i)(X(i)))^D$

(b) Suppose $\leq$ is a partial order making $(I, \leq)$ a net such that

(i) $X(i)$ converge weakly, and

(ii) $\{j : i \leq j\} \in D$ for each $i \in I$.

Then $X(i)$ converge weakly to X^D.

Proof: Write $X = X^D$.

(a) $\{X \geq r\} = \bigcap_n (F_n(i))^D$, where $F_n(i) = \{X(i) \geq r - \frac{1}{n}\}$, so X is $\mathcal{F}$-measureable. Let $I_{k,n} = [\frac{k}{n}, \frac{k+1}{n})$. Then

$$P(X) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{k}{n} P(X \in I_{k,n})$$
$$= \lim_{n\to\infty} \lim_{\varepsilon \downarrow 0} \sum_{k=0}^{n} \frac{k}{n} P((\, X(i) \in I_{k,n} - \varepsilon \,)^D)$$
$$= \lim_{n\to\infty} \lim_{\varepsilon \downarrow 0} \sum_{k=0}^{n} \frac{k}{n} (P(i)(X(i) \in I_{k,n} - \varepsilon))^D$$

$$= \lim_{n\to\infty} \lim_{\varepsilon \downarrow 0} \left(\sum_{k=0}^{n} \frac{k}{n} P(i)(X(i) \in I_{k,n} - \varepsilon) \right)^D \quad \text{(by Prop. 3.5)}$$

$$= (P(i)(X(i)))^D ,$$

since for $\varepsilon \leq \frac{1}{n}$, $\left| P(i)(X(i)) - \sum_{k=0}^{n} \frac{k}{n} P(i)(X(i) \in I_{k,n} - \varepsilon) \right| < \frac{2}{n}$.

(b) If $P(i)(\varphi(X(i))) \to r$, then for each $\varepsilon > 0$, $\{ i : |P(i)(\varphi(X(i))) - r| < \varepsilon \} \in D$, so $(P(i)(\varphi(X(i))))^D = r$. ■

Now we consider ultraproducts of graded probability spaces. But first note that ultraproducts commute with set products: that is, $(\Omega(i)^n)^D$ is in 1-1 correspondence with $((\Omega(i))^D)^n$ by

$$(\omega_1(i), \ldots, \omega_n(i))^D \longleftrightarrow ((\omega_1(i))^D, \ldots, (\omega_n(i))^D)$$

Henceforth we will identify the two sets via this correspondence.

3.9. Definition: If $(\Omega(i), \mathcal{F}_n(i), P_n(i))_{n\in\mathbb{N}}$ is a g.p.s. for each $i \in I$, then their ultraproduct with respect to the countably incomplete ultrafilter D on I is the g.p.s. $(\Omega, \mathcal{F}_n, P_n)$, where $(\Omega^n, \mathcal{F}_n, P_n)$ is the ultraproduct of $(\Omega(i)^n, \mathcal{F}_n(i), P_n(i))$ for each $n \in \mathbb{N}$. ■

This, too, makes sense:

3.10. Proposition (Keisler [8]): $(\Omega, \mathcal{F}_n, P_n)_{n\in\mathbb{N}}$ as defined in 3.9 is a g.p.s.

Proof: We already know that each $(\Omega^n, \mathcal{F}_n, P_n)$ is a probability space, so we just need to check the Fubini Property. It will suffice to show that the various clauses of the property hold of each $F \in (\mathcal{F}_n(i))^D$, since $(\mathcal{F}_n(i))^D$ is a field generating $\mathcal{F}_n$.

(i) If $F = (F(i))^D \in (\mathcal{F}_n(i))^D$, then

$$P_n(F) = (P_n(i)(F(i)))^D = (P(i)(F(i)^{\pi}))^D = P(F^{\pi})$$

(2) If $F \in (\mathcal{F}_{n+m}(i))^D$, $\vec{\omega} = (\vec{\omega}(i))^D \in \Omega^m$, then $F(\cdot, \vec{\omega}) = (F(i)(\cdot, \vec{\omega}(i)))^D \in (\mathcal{F}_n(i))^D$, since $F(i)(\cdot, \vec{\omega}(i)) \in \mathcal{F}_n(i)$ for each $i \in I$.

(3) $\{\vec{\omega} \in \Omega^m : P_n(F(\cdot, \vec{\omega})) \geq r\} = \bigcap_k (G_k(i))^D \in \mathcal{F}_m$, where

$$G_k(i) = \{\vec{\omega}(i) \in \Omega(i)^m : P_n(i)(F(i)(\cdot, \vec{\omega}(i))) > r - \frac{1}{k})\}$$

(4)
$$P_{n+m}(F) = (P_{n+m}(i)(F(i)))^D = (\int P_n(i)(F(i)(\cdot, \vec{\omega}(i))(P_m(i)(\vec{\omega}(i)))^D$$
$$= \int P_n(F(\cdot, \vec{\omega}))dP_m(\omega)$$

by two applications of Prop. 3.8(a). ■

Now we want to prove a result like Prop. 3.8(b) but for arrays. It will follow easily from the following.

3.11. Definition: Let $X : \Omega^n \to [0,1]$ be a relation on a g.p.s. $(\Omega, \mathcal{F}_n, P_n)$. For $\sigma \in \mathbb{N}^n$, $k \in \mathbb{N}$, define $X^{ED}(\sigma, k)$ ($X^{DRE}(\sigma, k)$) on $(\Omega_{n\times k}, \mathcal{F}_{n\times k}, P_{n\times k})$ (resp. $(\Omega_k, \mathcal{F}_k, P_k)$) by

$$X^{ED}(\sigma, k)(\vec{\omega}) = X(\omega(1, \sigma(1)), \ldots, \omega(n, \sigma(n)))$$

(resp. $X^{DRE}(\sigma, k)(\vec{\omega}) = X(\omega(\sigma(1)), \ldots, \omega(\sigma(n)))$. ■

3.12. Proposition: Let $\sigma_1, \ldots, \sigma_m \in k^n$.

(a) $(X^{ED}_{\sigma_1}, \ldots, X^{ED}_{\sigma_m}) \sim (X^{ED}(\sigma_1, k), \ldots, X^{ED}(\sigma_m, k))$.

(b) The same holds with DRE replacing ED.

Proof: Immediate from the definitions and the Fubini Property. ■

3.13. Proposition: Let D be a countably incomplete ultrafilter on I, $X(i) : \Omega(i)^n \to [0,1]$ relations on graded probability spaces $(\Omega(i), \mathcal{F}_m(i), P_m(i))_{m\in\mathbb{N}}$, $i \in I$, $X = X^D$ the ultraproduct of $X(i)$, $i \in I$, on $(\Omega, \mathcal{F}_m, P_m)_{m\in\mathbb{N}}$ the ultraproduct of $(\Omega(i), \mathcal{F}_m(i), P_m(i))_{m\in\mathbb{N}}$, $i \in I$.

(a) For each $k \in \mathbb{N}$, $\sigma \in k^n$, $X^{ED}(\sigma, k) = (X(i)^{ED}(\sigma, k))^D$, and $X^{DRE}(\sigma, k) = (X(i)^{DRE}(\sigma, k))^D$.

(b) If $(I, \leq)$ is a net and $\{j : i \leq j\} \in D$ for each $i \in I$, then if $\{X(i)_\sigma^{ED}\}_{\sigma\in\mathbb{N}^n}$ converges weakly, it converges weakly to $\{X_\sigma^{ED}\}_{\sigma\in\mathbb{N}^n}$ (the same holds with DRE instead of ED).

Proof: (a) is immediate. To prove (b), we need only, by Prop. 3.12, to show that for each $\varphi \in C([0,1]^n)$, $k \in \mathbb{N}$, $\sigma_1, \ldots, \sigma_m \in k^n$,
$\varphi(X(i)^{ED}(\sigma_1, k), \ldots, X(i)^{ED}(\sigma_m, k)) \to \varphi(X^{ED}(\sigma_1, k), \ldots, X^{ED}(\sigma_m, k))$
(or the same thing with RCE instead of ED). But this is immediate by (a) and Prop. 3.12. ∎

4. Representation by Sampling

The main result of this section is that every DRE array, and hence every ED or DWE array has the same distribution as an array induced by sampling a relation on a g. p. s. We will state the theorem for countable families of arrays; by adding some slight further considerations, this could be modified to give the same result for arrays taking values in any complete separable metric space. With a slight change in the proof, we could also show that the theorem holds for uncountable families of arrays.

The theorem appears in [7], cloaked in different terminology.

4.1. Theorem (Sampling Representation Theorem): (a) The family of arrays $\{Y^k_\sigma : \sigma \in \mathbb{N}^{n(k)}, k, n(k) \in \mathbb{N}\}$ is DRE iff $\{Y^k_\sigma\} \sim \{(X^k)^{DRE}_\sigma\}$ where $X^k : \Omega^{n(k)} \to [0,1]$ are relations on a g. p. s. $(\Omega, \mathcal{F}_n, P_n)_{n \in \mathbb{N}}$.

(b) $\{Y^k_s : s \in C(\mathbb{N}, n(k)), k \in \mathbb{N}\}$ is DWE iff $\{Y^k_s\} \sim \{X^k_s\}$ for some family of relations X^k, $k \in \mathbb{N}$, such that $(X^k)^\pi = X^k$ for all $\pi \in n(k)!$, $k \in \mathbb{N}$.

Proof: In (b) let $\{Y^k_\sigma\}$ be the DRE arrays induced by $\{Y^k_s\}$ by $Y^k_\sigma = Y^k_s$ where $s = \{\sigma(i) : i \leq n(k)\}$. If $\{(X^k)^{DRE}_\sigma\} \sim \{Y^k_\sigma\}$ then clearly $(\overline{X}^k)^\pi = \overline{X}^k$ a. s. for $\pi \in n(k)!$, $k \in \mathbb{N}$. So if we just let $X^k = \min\{(\overline{X}^k)^\pi : \pi \in n(k)!\}$ for each $k \in \mathbb{N}$, then $\{X^k_s\} \sim \{Y^k_s\}$. Hence (b) will follow from (a).

(a) Let $Y^k_\sigma : \overline{\Omega} \to [0,1]$, $\sigma \in \mathbb{N}^{n(k)}$, $k \in \mathbb{N}$, $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$ a probability space. The "if" part is immediate by the Fubini Property, as was remarked in the introduction. To prove the "only if", let $\mathcal{E}$, $\mathcal{E}_m$ $m \in \mathbb{N}$, be as in 2.3. The DRE version of the Hewitt-Savage 0-1 Law and a routine reverse martingale argument show that if $\varphi \in C([0,1]^\ell)$, $k_1, \ldots, k_\ell \in \mathbb{N}$, $\sigma_i \in \mathbb{N}^{n(k_i)}$, $1 \leq i \leq \ell$, then

$$(*)\qquad \overline{P}(\varphi(Y_{\sigma_1}^{k_1},\dots,Y_{\sigma_\ell}^{k_\ell})) = \overline{P}(\varphi(Y_{\sigma_1}^{k_1},\dots,Y_{\sigma_\ell}^{k_\ell})\,|\,\mathfrak{E})\ \text{a.s.}$$

$$= \lim_m \overline{P}(\varphi(Y_{\sigma_1}^{k_1},\dots,Y_{\sigma_\ell}^{k_\ell})\,|\,\mathfrak{E}_m)\ \text{a.s.}$$

$$= \lim_m \frac{1}{m!}\sum_{\pi\in m!} \varphi(Y_{\sigma_1\pi}^{k_1},\dots,Y_{\sigma_\ell\pi}^{k_\ell})\ \text{a.s.}$$

Choose $\omega \in \overline{\Omega}$ such that this equality holds of $\varphi(Y_{\sigma_1}^{k_1},\dots,Y_{\sigma_\ell}^{k_\ell})$ for each $\sigma_i \in \mathbb{N}^{n(k_i)}$, $k_1,\dots,k_\ell, \ell \in \mathbb{N}$, and φ in a countable dense subset D_ℓ of $C([0,1]^\ell)$. For $m \in \mathbb{N}$, define a g.p.s. $(\Omega(m), \mathcal{F}(m)_n, P(m)_n)_{n\in\mathbb{N}}$ and relations $X^k(m)$ on $\Omega^{n(k)}$ by: $\Omega(m) = m$, $\mathcal{F}(m)_n = \mathcal{P}(m^n)$, the set of all subsets of m^n, and let $P(m)_n$ be the uniform probability on m^n;

$$X^k(m)(i_1,\dots,i_{n(k)}) = Y^k_{(i_1,\dots,i_{n(k)})}(\omega),\quad i_1,\dots,i_{n(k)} \le m\ .$$

We show that for each $k_1,\dots,k_\ell$, $\sigma_i \in \mathbb{N}^{n(k)}$, $\ell \in \mathbb{N}$, $\varphi \in D_\ell$, $k = \max(k_i)$,

$$P(m)(X(m)) = P(m)(\varphi((X^{k_1}(m))^{DRE}(\sigma_1,k),\dots,(X^{k_\ell}(m))^{DRE}(\sigma_\ell,k)))$$

$$\to \overline{P}(\varphi(Y_{\sigma_1}^{k_1},\dots,Y_{\sigma_\ell}^{k_\ell})) = \overline{P}(Y_{\sigma_1,\dots,\sigma_\ell})$$

as $m \to \infty$, where $X(m)$ and $Y_{\sigma_1,\dots,\sigma_\ell}$ are introduced as abbreviations. For

$$P(m)(X(m)) = \frac{1}{m^k}\sum_{\vec{p}\in m^k} X(m)(\vec{p})$$

$$\to \lim_m \frac{(m-k)!}{m!}\sum_{\vec{p}\in P(m,k)} X(m)(\vec{p})$$

$$= \lim_m \frac{1}{m!}\sum_{\pi\in m!} Y_{\sigma_1^\pi,\dots,\sigma_\ell^\pi}(\omega)$$

$$= \bar{P}(Y_{\sigma_1, \ldots, \sigma_\ell})$$

by (*). Thus $\{(X(m)^k)_\sigma^{DRE}\}$ converges weakly to $\{Y_\sigma^k\}$. So if D is a countably incomplete ultrafilter on $\mathbb{N}$, $(\Omega, \mathcal{F}_n, P_n)_{n \in \mathbb{N}}$ the ultraproduct of $(\Omega(m), \mathcal{F}(m)_n, P(m)_n)$ with respect to D, and $X^k = (X^k(m))^D$, then it follows easily by Props. 3.12, 3.13(a), and 3.8(b) that $\{(X^k)_\sigma^{DRE}\} \sim \{X_\sigma^k\}$. ∎

4.2. Remark: This is probably a good place to discuss a couple of facts about arrays, as they may be compared with the corresponding facts about relations.

(1) DRE arrays: Let $\{X_\sigma\}_{\sigma \in \mathbb{N}^n}$ be DRE. If $\sigma \in \mathbb{N}^n$ has $\sigma(i) = \sigma(j)$ for some $i < j \leq n$, then obviously there is no $\pi \in \mathbb{N}$ such that $\sigma^\pi(i) \neq \sigma^\pi(j)$ hence $\{X_\sigma : \sigma(i) = \sigma(j)\}$ bears no special relation to $\{X_\sigma : \sigma(i) \neq \sigma(j)\}$ beyond being jointly dissociated. In fact we can dissect X into a family of arrays

$$\{Y^\sim : \sim \text{ an equivalence relation on } n,\ \sigma \in P(\mathbb{N}, n)\}$$

where $Y_\sigma^\sim = Y_{\hat{\sigma}}$, $\hat{\sigma}(i) = \sigma(j)$ where j is the least number such that $i \sim j$, and any such DRE family $\{Y_\sigma^\sim\}$ can be used to construct an $\{X_\sigma\}$ in this way. This is related to the following fact about relations: suppose $X : \Omega^n \to [0,1]$ is a relation on a g.p.s. $(\Omega, \mathcal{F}_n, P_n)_{n \in \mathbb{N}}$ and $P_1(\omega) = 0$, for every $\omega \in \Omega$. Then if $D = \{\vec{\omega} \in \Omega^n : \omega(i) = \omega(j)$ for some $i < j\}$, $P_n(D) = 0$. Hence we can change X on D (and hence change $\{X_\sigma : \sigma(i) = \sigma(j)$ for some $i < j\}$) however we like, without changing the distribution of $\{X_\sigma^{DRE}\}_{\sigma \in P(\mathbb{N}, n)}$. Since it is inconvenient to deal with an object affected by change X on a set of measure zero, we will henceforth consider only $\{X_\sigma^{DRE}\}_{\sigma \in P(\mathbb{N}, n)}$ as the DRE array induced by X, and we will generally consider DRE arrays $\{X_\sigma\}$ as defined only for $\sigma \in P(\mathbb{N}, n)$, unless specified otherwise. With this restriction, $X = Y$ a.s. implies $\{X_\sigma^{DRE}\} \sim \{Y_\sigma^{DRE}\}$.

(2) ED arrays are a special case of RCE arrays, so Theorem 4.1 implies that any ED array has the same distribution as $\{X_\sigma^{DRE}\}_{\sigma \in \mathbb{N}^n}$ for some relation X. Let us consider this situation a little further.

(i) If $X : \Omega^n \to [0,1]$ is a relation on $(\Omega, \mathcal{F}_n, P_n)_{n \in \mathbb{N}}$, then

$$\{X_\sigma^{ED}\}_{\sigma \in \mathbb{N}^n} \sim \{Y_\sigma^{DRE}\}_{\sigma \in \mathbb{N}^n} ,$$

where Y on $(\Omega^n, \mathcal{F}_{nm}, P_{nm})_{m \in \mathbb{N}}$ is defined by

$$Y(\omega_1, \ldots, \omega_n) = X(\omega_1(1), \ldots, \omega_n(n)) \qquad \omega_i \in \Omega^n .$$

(ii) If $\{X_\sigma^{DRE}\}_{\sigma \in \mathbb{N}^n}$ is actually ED, then

$$\{X_\sigma^{ED}\} \sim \{X_\sigma^{DRE}\}$$

This follows easily using Prop. 3.12.

(iii) For $\sigma \in \mathbb{N}^n$, define $f(\sigma)$ by

$$f(\sigma)(i) = \sigma(i)n + i$$

Then if $Y_\sigma = X_{f(\sigma)}^{DRE}$, $\sigma \in \mathbb{N}^n$, then

$$\{Y_\sigma\} \sim \{X_\sigma^{ED}\}$$

Thus the ED array induced by X is in a sense contained in the induced DRE array. ■

5. Results on σ-Fields.

This section contains technical results about the relationships between a relation X on a g.p.s. $(\Omega, \mathcal{F}_n, P_n)$, and various σ-fields on $\Omega^{\mathbb{N}\times n}$ which are defined either from $\mathcal{F}_m$ $m \in \mathbb{N}$, or from $\{X_\sigma^{ED}\}_{\sigma\in\mathbb{N}^n}$. Most of these results state that some σ-field contains all the information (in a non-technical, heuristic sense) about X, or $\{X_\sigma^{ED}\}$ that some other σ-field does; the main result is that all the information about X contained in a certain σ-field defined from $(\mathcal{F}_n : m \in \mathbb{N})$ is contained in a σ-field of similar form defined from X. This and its ancillary results are the main tools in the analysis of relations and arrays carried out in section 7.

In this section we shall be mainly concerned with ED arrays (because they contain less information about the relations that induce them than do DRE arrays). Hence, in this section only, if X is a relation, shall we write $\{X_\sigma\}$ for $\{X_\sigma^{ED}\}$.

The formal notion of conditional independence embodies the heuristic notion of "containing all the information".

5.1. Definition: (1) Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G}, \mathcal{H}_i$ $i \in I$ sub-σ-algebras of $\mathcal{F}$. We say that $\mathcal{H}_i$, $i \in I$, are c.i. given $\mathcal{G}$, if I_0 is a finite subset of I, $H_i \in \mathcal{H}_i$, $i \in I_0$,

$$P(\prod_{i\in I_0} H_i | \mathcal{G}) = \prod_{i\in I_0} P(H_i | \mathcal{G}) .$$

In the case where $I = \{1, 2\}$, a simple computation shows that this is the same as saying

$$P(H|\mathcal{H}_2, \mathcal{G}) = P(H|\mathcal{G})$$

for every $H \in \mathcal{H}_1$; in this case we will often use the terminology "$\mathcal{H}_1$ is independent of $\mathcal{H}_2$ given $\mathcal{G}$", as if the relation were not symmetric.

Of course, we extend the definition to random variables in the obvious way.

(2) Let $(\Omega, \mathcal{F}_n, P_n)_{n\in\mathbb{N}}$ be a g.p.s.

(3) A relation on Ω^n, $\Omega^{\mathbb{N}}$, or $\Omega^{\mathbb{N}\times n}$ is supported by, or has support in $t \subseteq n$ if $X(\vec{\omega}) = X(\vec{\omega}')$ whenever

$$(*) \qquad \vec{\omega}(i) = \vec{\omega}'(i) , \qquad i \in t.$$

We will identify any such X with X' on Ω^s, $t \subseteq s$ such that $X(\vec{\omega}) = X'(\vec{\omega}')$, for $\vec{\omega}, \vec{\omega}'$ satisfying (*).

(4) If $s \subseteq n$ (or $\mathbb{N}$, or $\mathbb{N} \times n$) identify $\mathcal{F}_s$ with the class of $\mathcal{F}_n$ (or $\mathcal{F}_{\mathbb{N}}$ or $\mathcal{F}_{\mathbb{N}\times n}$)- measureable subsets of Ω^n (etc.), which have support in s. For s a set of subsets of n, etc., let

$$\mathcal{F}_S = \bigvee_s \mathcal{F}_s .$$

(5) For X a relation on Ω^n, S a set of subsets of $\mathbb{N} \times n$ let $\mathcal{G}_S^X$ denote the σ-algebra

$$\bigcap_m (\mathcal{F}(\{X_\sigma\}) \cap \mathcal{F}_{S(m)}) = \bigcap_m \mathcal{F}(X_\sigma : \exists s \in S(m)(\{(\sigma(j, i) : i \leq n\} \subseteq s)) ,$$

where $S(m) = \{s \cup (\mathbb{N}-m) \times n : s \in S\}$. If $S = \{s\}$, we write $\mathcal{G}_s^X$ for $\mathcal{G}_S^X$.

(6) Call $\bigcup S = \bigcup_{s\in S} s$ the support of S . ■

5.2. Example: (a) $\mathcal{F}_{C(n,1)} = (\mathcal{F}_1)^n$.

(b) If $X = C(\mathbb{N} \times n,\ n-1)$, $\mathcal{F}_S$ is "all the n-1 dimensional information", and is $\mathcal{G}^X_S$ "all the n-1 dimensional information which $\{X_\sigma\}$ knows about itself". For $n = 2$ this $\mathcal{G}^X_S$ is the same as the "shell" σ-algebra of Aldous[1].

5.3. Remark: Clearly if

$$\{u : \exists\, s \in S \text{ s.t. } u \subseteq s)\} \subseteq \{u : \exists t \in T \text{ s.t. } u \subseteq t\}$$

then $\mathcal{G}^X_S \subseteq \mathcal{G}^X_T$. ■

In the following we will write $\mathcal{G}$ for $\mathcal{F}(\{X_\sigma\})$. $\mathcal{G}_S$ for $\mathcal{G}^X_S$. Let $\mathcal{N}$ denote the nullsets of $P_{\mathbb{N}\times n}$.

5.4. Lemma: Any $G \in \mathcal{G}_S$ is fixed by any finite, strict permutation which fixes the support of S.

Proof: Suppose π is a finite strict permutation which fixes $\bigcup S$ and any element outside $m \times n$. By definition, any $G \in \mathcal{G}_S$ has support in $\bigcup S \cup (\mathbb{N}-m) \times n$; clearly π fixes any set which has support in the subset of $\mathbb{N} \times n$ fixed by π. ■

5.5. Proposition: (a) For each S of finite support, $\mathcal{G}_S$ is essentially contained in $\mathcal{F}_S$, i.e.,

$$\mathcal{G}_S \subseteq \mathcal{F}_S \vee \mathcal{N}.$$

(b) If $t \subseteq \mathbb{N} \times n$ is finite, $\mathcal{G}_S$ (resp. $\mathcal{F}_S$) is c.i. of $\mathcal{F}_t$ given $\mathcal{G}_{S\cap t}$ (resp. $\mathcal{F}_{S\cap t}$) where $S\cap t = \{s \cap t : s \in S\}$.

(c) (a) and (b) hold without finiteness conditions on $\bigcup S$ or t.

Proof: (a) Let $G \in \mathcal{G}_S$. We will show that for each $\varepsilon > 0$ there is an $\mathcal{F}_S$-measureable function Z such that $P(|Z-G|) < \varepsilon$.

Let $\cup S \subseteq m \times n$, and let Y be a bounded $\mathcal{G} \cap \mathcal{F}_{S(m)}$-measureable function (where $S(m)$ is as in 5.1(5)) such that $P(|Y-G|) < \varepsilon$. We may take Y to be of finite support t, since $\mathcal{G} \cap \mathcal{F}_{S(m)}$ is generated by sets of finite support. Let us write

$$Y(\vec{\omega}) = Y(\vec{\omega}_1, \vec{\omega}_2)$$

$\vec{\omega}_1 \in \Omega^{m \times n}$, $\vec{\omega}_2 \in \Omega^{t-(m \times n)}$. Let π_i ; $i \in \mathbb{N}$ be a sequence of strict finite permutations of $\mathbb{N} \times n$, fixing $m \times n$, but such that, for $i \neq j$,

$$\pi_i(t-(m \times n)) \cap \pi_j(t-m \times n) = \phi$$

Then for each fixed $\vec{\omega}_1$, the sequence $Y(\vec{\omega}_1, \vec{\omega}_2)^{\pi_i} = Y(\vec{\omega}_1, \vec{\omega}_2^{\pi_i})$, $i \in \mathbb{N}$, is i.i.d. (by the Fubini Property), and so by the weak law of large numbers, $\frac{1}{N} \sum_{i \leq N} Y^{\pi_i}$ converges in probability, say to $Z(\vec{\omega}_1)$, which we may take to be bounded.

$$P(|Z-G|) \leq \lim_N P(\frac{1}{N} \sum_{i \leq N} |Y^{\pi_i}-G|) = P(|Y-G|) < \varepsilon,$$

(the equality since G is fixed by each π_i). So it will suffice to show that Z is essentially $\mathcal{F}_S$ measureable. Now since

$$\frac{1}{N} \sum_{i \leq N} Y(\vec{\omega}_1, \vec{\omega}_2)^{\pi_i} \to Z(\vec{\omega}_1) \qquad \text{a.s. } (\vec{\omega}_1, \vec{\omega}_2),$$

there is, by the Fubini property, a fixed $\vec{\omega}_0 \in \Omega^{(\mathbb{N}-m)\times n}$ such that

$$\frac{1}{N} \sum_{i \leq N} Y(\vec{\omega}_1, \vec{\omega}_0^{\pi_i}) \to Z(\vec{\omega}_1) \qquad \text{a.s. } (\vec{\omega}_1) .$$

But each $Y(\vec{\omega}_1, \vec{\omega}_0^{\pi_i})$ is $\mathcal{F}_S$-measureable.

(b) The idea is basically the same as in (a). The two cases for $\mathcal{A}_S$ and $\mathcal{F}_S$ are similar, so we will do only the former. If $t \subseteq m \times n$, it will suffice to show that for every $G \in \mathcal{A} \cap \mathcal{F}_{S(m)}$, $P(G|\mathcal{F}_t)$ is $\mathcal{A}_{S \cap t} \vee \mathcal{N}$ measureable, since it follows by (a) that $\mathcal{A}_{S \cap t}$ is essentially contained in $\mathcal{F}_{S \cap t} \subseteq \mathcal{F}_t$. We may take G to be of finite support, as $\mathcal{A} \cap \mathcal{F}_{S(m)}$ is generated by such sets. Write $G(\vec{\omega}) = G(\vec{\omega}_1, \vec{\omega}_2)$ $\vec{\omega}_1 \in \Omega^t$, $\omega_2 \in \Omega^{u-t}$ where u is a finite support of G. As in (a), let π_i, $i \in \mathbb{N}$ be strict finite permutations of $\mathbb{N} \times n$ fixing t, such that $G(\vec{\omega}_1, \vec{\omega}_2)^{\pi_i} = G(\vec{\omega}_1, \vec{\omega}_2^{\pi_i})$, $i \in \mathbb{N}$, are i.i.d. for each fixed ω_1; assume also that $\pi_i(u-t) \subseteq (\mathbb{N}-i) \times n$. Then by the strong law of large numbers,

$$\lim_N \frac{1}{N} \sum_{i \leq n} G(\vec{\omega})^{\pi_i} = P(G|\mathcal{F}_t) \text{ a.s. } .$$

the left-hand side, however, is $\mathcal{A}_{S \cap t}$-measureable.

(c) The point is then we can now get our hands on a fact which seems obvious but is not quite, namely that $\mathcal{A}_S$ is essentially generated by those of its members which essentially have finite support (this follows directly from the definitions for $\mathcal{F}_S$). For if $G \in \mathcal{A}_S$,

$$G = P(G|\mathcal{F}_{\mathbb{N}\times n}) = \lim_m P(G|\mathcal{F}_{m\times n}) \text{ a.s.}$$

by the martingale convergence theorem. By (b), $P(G|\mathcal{F}_{m\times n})$ may be taken $\mathcal{G}_{S\cap m\times n}$-measureable. Thus

$$(*) \qquad \mathcal{G}_S \subseteq \bigvee_m \mathcal{G}_{S\cap m\times n} \vee \mathcal{N}.$$

Hence

$$G \in \bigvee_m \mathcal{G}_{S\cap m\times n} \vee \mathcal{N} \subseteq \bigvee_m \mathcal{F}_{S\cap m\times n} \vee \mathcal{N} = \mathcal{F}_S \vee \mathcal{N}$$

the set containment following by (a). And

$$\begin{aligned} P(G|\mathcal{F}_t) &= \lim_m P(G|\mathcal{F}_{t\cap m\times n}) \\ &= \lim_m P(G|\mathcal{G}_{S\cap t\cap m\times n}) \\ &= \lim_m P(G|\mathcal{G}_{S\cap t}) \end{aligned}$$

by (*). The same holds with $\mathcal{F}_S$ replacing $\mathcal{G}_S$. ■

5.6. Corollary: (a) $\mathcal{F}_t$ is c.i. of $\mathcal{G}_S$ given $\mathcal{G}_{S\cap t}$ (resp. $\mathcal{F}_S$ given $\mathcal{F}_{S\cap t}$).

(b) If for each $i \neq j \in I$, $t_i \cap t_j \subseteq s$ for some $s \in S$, then $\mathcal{G}_{t_i}$ $i \in I$ (resp. $\mathcal{F}_{t_i}$, $i \in I$), are c.i. given $\mathcal{G}_S$ ($\mathcal{F}_S$).

(c) If X and Y are relations, $S = \{s \subseteq \mathbb{N} \times n;\ |s| < n,\ |s \cap \mathbb{N} \times \{i\}| \leq 1,\ i \leq n\}$,

$$\{\mathcal{D}(X_\sigma \mid \mathcal{G}_S^X)\}_{\sigma\in\mathbb{N}^n} \sim \{\mathcal{D}(Y_\sigma \mid \mathcal{G}_S^Y)\}_{\sigma\in\mathbb{N}^n},$$

then $\{X_\sigma\} \sim \{Y_\sigma\}$ (same with $\mathcal{F}$ replacing $\mathcal{G}^X$ and $\mathcal{G}^Y$).

Proof: We will use $\mathcal{G}$ throughout, as the proofs are the same with $\mathcal{F}$ replacing $\mathcal{G}$.

(a) Immediate from Prop. 5.5.

(b) For $i \in I$, let Y_i be $\mathcal{G}_{t_i}$-measureable, and let $T_i = \{t_j : j \neq i\}$. Certainly Y_i is c.i. of Y_j, $j \neq i$, given $\mathcal{G}_{S\cup T_i}$. But

$$\mathcal{D}(Y_i \mid \mathcal{G}_{S\cup T_i}) = \mathcal{D}(Y_i \mid \mathcal{G}_{(S\cup T_i)\cap t_i}) \text{ a.s.}$$

By 5.5(b), (c). But since every $t_i \cap t_j$, $i \neq j$, is contained in a member of S, $\mathcal{G}_{(S\cup T_i)\cap t_i} = \mathcal{G}_{S\cap t_i}$. But by 5.5, again,

$$\mathcal{D}(Y_i \mid \mathcal{G}_{S\cap t_i}) = \mathcal{D}(Y_i \mid \mathcal{G}_S) \text{ a.s.}$$

(c) The distributions of $\{X_\sigma\}$ and $\{Y_\sigma\}$ are determined by the distributions of their conditional distributions: for if $\varphi_1, \dots, \varphi_m \in C([0,1])$, then

$$P(\varphi_1(X_{\sigma_1})\dots\varphi_m(X_{\sigma_m})) = P(P(\varphi_1(X_{\sigma_1})\dots\varphi_m(X_{\sigma_m}) \mid \mathcal{G}_S))$$

$$= P(\prod_{i\le m} P(\varphi_i(X_{\sigma_i}) \mid \mathcal{G}_S)) \qquad \text{(by (b))}$$

$$= P(\prod_{i\le m} \mathcal{D}(X_{\sigma_i} \mid \mathcal{G}_S)(\varphi_i)) . \quad \blacksquare$$

Our aim now is to prove that any $G \in \mathscr{G}$ is c.i. of $\mathscr{F}_S$ given $\mathscr{G}_S$, for S any set of subsets of $\mathbb{N} \times n$. We will begin by proving by induction a case where S is finite. In this induction the following lemma will provide the induction step.

Let $\vec{1}$ denote the sequence $(1, \ldots, 1) \in \mathbb{N}^n$.

5.7. Lemma: Let X, F be relations on Ω^n such that F has support in $s \subseteq n$. Then if S is a set of subsets of $1 \times n$ such that $1 \times s \in S$, then $X_{\vec{1}}$ is c.i. of $\mathscr{G}^{X,F}_{S(1)}$ given $\mathscr{G}_S = \mathscr{G}^X_S$, where S(1) is as in 5.1(4).

Proof: Let us write X for $X_{\vec{1}}$, for simplicity. Since $\mathscr{G}^{X,F}_{S(1)}$ is generated by the semi-field of sets of the form GH, where $G \in \mathscr{G} \cap \mathscr{F}_{S(1)}$, $H \in \mathscr{G}^F$ (for since $\mathbb{N} \times s \in S(1)$, $\mathscr{G}^F = \mathscr{G}^F_{S(1)}$) both have finite support, it will suffice to show that

$$P(\varphi(X)GH) = P(P(\varphi(X)\,|\,\mathscr{G}_S)GH) \quad \text{for such } G,\ H \text{ and } \varphi \in C([0,1]).$$

Let G, H have support in $m \times n$. Now let π_i be finite restricted permutations of $\mathbb{N} \times n$ fixing $\mathbb{N} \times s$, such that $\pi_i(m \times (n-s))$, $i \in \mathbb{N}$, are disjoint. It follows by Corollary 5.6(b) that $(\varphi(X)G)^{\pi_i}$, $i \in I$, are c.i. given $\mathscr{G}_{\mathbb{N} \times s}$. Now for $N \in \mathbb{N}$

$$P(\varphi(X)GH) = \frac{1}{N} \sum_{i \leq N} P((\varphi(X)GH)^{\pi_i})$$

$$= \frac{1}{N} \sum_{i \leq N} P((\varphi(X)G)^{\pi_i} H),$$

the first equality by the Fubini Property, the second since π_i, $i \in I$, fix H. But since $(\varphi(X)G)^{\pi_i}$ are c.i. given $\mathcal{G}_{\mathbb{N}\times s}$, it follows by an application of Chebyshev's Inequality, conditional on $\mathcal{G}_{\mathbb{N}\times s}$, that

$$\frac{1}{N}\sum_{i\leq N}[(\varphi(X)G)^{\pi_i} - P((\varphi(X)G)^{\pi_i}|\mathcal{G}_{\mathbb{N}\times s})]$$

converges to zero in probability. Since each π_i fixes $\mathcal{G}_{\mathbb{N}\times s}$, $P((\varphi(X)G^{\pi_i}|\mathcal{G}_{\mathbb{N}\times s}) = P(\varphi(X)G|\mathcal{G}_{\mathbb{N}\times s})^{\pi_i}$, $i \in I$. Hence, since everything is bounded,

$$P(\frac{1}{N}\sum_{i\leq N}[(\varphi(X)G)^{\pi_i} - P((\varphi(X)G)^{\pi_i}|\mathcal{G}_{\mathbb{N}\times s})]H)$$

$$= \frac{1}{N}\sum_{i\leq N}P((\varphi(X)G)^{\pi_i}H) - \frac{1}{N}\sum_{i\leq N}P(P(\varphi(X)G|\mathcal{G}_{\mathbb{N}\times s})^{\pi_i}H)$$

$$= P(\varphi(X)GH) - P(P(\varphi(X)G|\mathcal{G}_{\mathbb{N}\times s})H)$$

tends to zero. Being independent of N, it is zero. But

$$P(P(\varphi(X)G|\mathcal{G}_{\mathbb{N}\times s})H) = P(P(P(\varphi(X)G|\mathcal{G}_{S(1)})|\mathcal{F}_{\mathbb{N}\times s})H)$$

by 5.5(b), (c)

$$= P(P(\varphi(X)|\mathcal{G}_{S(1)})GH)$$

$$= P(P(\varphi(X)|\mathcal{G}_S)GH)$$

by Prop. 5.5(b). ∎

5.8. Theorem: For S a set of subsets of $\mathbb{N}\times n$, $X = X_{\vec{I}}$ is c.i. of $\mathcal{F}_S$ given $\mathcal{G}_S$.

Proof: By Prop. 5.5(b), we may as well take S to be a set of subsets of $1 \times n$, and therefore finite. We will prove by induction on $|T|$, that for each $T \subseteq S$, any Y with support in $1 \times n$ is c.i. of $\mathcal{F}_T$ given $\mathcal{G}_S^Y$. This is trivial if T is empty. Suppose then that it holds for $T \subseteq S$, $s \in S - T$, and we will prove it for $T \cup \{s\}$. Since $\mathcal{F}_{T \cup \{s\}}$ is generated by the semi-field of sets of the form FF', $F \in \mathcal{F}_s$, $F' \in \mathcal{F}_T$, it suffices to show

$$P(\varphi(Y)FF') = P(P(\varphi(Y) \mid \mathcal{G}_S^Y)FF')$$

for such F, F', and $\varphi \in C[0,1]$. But

$$P(\varphi(Y)FF') = P(P(\varphi(Y)F) \mid \mathcal{G}_S^{Y,F})F')$$

by induction hypothesis

$$= P(P(\varphi(Y) \mid \mathcal{G}_S^{Y,F})FF')$$

$$= P(P(\varphi(Y) \mid \mathcal{G}_S^{Y})FF')$$

by Lemma 5.7. ∎

5.9. Corollary: $\mathcal{G}$, is c.i. of $\mathcal{F}_S$ given $\mathcal{G}_S$, S any set of subsets of $\mathbb{N} \times n$.

Proof: We show that $G \in \mathcal{G}$ of finite support, say in $t \subseteq \mathbb{N} \times n$, is independent of $\mathcal{F}_S$ given $\mathcal{G}_S$. By Prop. 5.5(b) we can take $\cup S \subseteq t$. Let $t = \{t_1, \ldots, t_m\}$. Induce G' on Ω^m by

$$G'(\omega_1, \ldots, \omega_m) = G(\vec{\omega}'),$$

where $\vec{\omega}' \in \Omega^{\mathbb{N}\times n}$, $\vec{\omega}'(t_i) = \omega_i$, $i \leq m$. Let $g : \mathbb{N} \times m \to \mathbb{N} \times n$ be a bijection such that $g(1, i) = t_i$, $i \leq m$, and $g(i, j)$ has the same first coordinate as $g(i, k)$ for $i, j, k \in \mathbb{N}$. Then $G = (G')^g$, and it follows easily from the definitions that $(\mathcal{G}(G')_{S^{g-1}})^g \subseteq \mathcal{G}_S$, and $(\mathcal{F}_{S^{g-1}})^g = \mathcal{F}_S$. So it follows by the theorem that G is c.i. of $\mathcal{F}_S$ given $\mathcal{G}(G'_{S^{g-1}})^g$, hence given $\mathcal{G}_S$. ∎

The reader will no doubt anticipate that the foregoing analysis must somehow apply to general arrays, and not just to those induced by a relation on a g.p.s. So if for any ED or DRE array, or for RCE or RE arrays we define

$$\mathcal{G}_S^X \mathcal{G}_S(\{X_\sigma\}) = \bigcap_m \mathcal{F}(X_\sigma \quad : \exists s \in S \text{ such that } \forall i \leq n \quad (\sigma(i), i) \in s \cup (\mathbb{N}-m)\times n)$$

then by the Sampling Representation Theorem we easily get the following result:

5.10. Corollary: (a) If $\{(X, Z)_\sigma\}$ is ED (or DRE) then for S a set of subsets of $\mathbb{N} \times n$ $(\mathbb{N})$, X is c.i. of $\mathcal{G}_S^{X,Z}$ given $\mathcal{G}_S^X$.

(b) Cor. 5.6(b)(c) hold in general for ED arrays.

(c) (a) and (b) hold even for RCE (and in (a), RE) arrays.

Proof: (b) and the ED case of (a) are immediate from results for arrays induced by sampling. To get the DRE case of (a), consider the ED array $\{Y_\sigma\} = \{Y(\pi)_\sigma : \pi \in n!\}$, given by

$$Y(\pi)_\sigma = X_{\hat{\sigma}(\pi)}$$

where

$$\hat{\sigma}(\pi)(i) = (\sigma(\pi(i))-1)n + \pi(i)$$

Then $\mathcal{G}^X_S = \mathcal{G}^Y_T$, where

$$T = \{t \subseteq \mathbb{N} \times n : t = \{(k,i) : \exists i (1 \le i \le n \text{ and } kn+i \in s) \text{ for some } s \in S\}.$$

The result now follows by the ED case.

(c) Suppose for simplicity that $\{X_\sigma\}_{\sigma \in \mathbb{N}^n}$ is RCE. Define an ED array $\{\bar{X}_\tau\}_{\tau \in \mathbb{N}^{n+1}}$ so that $\{\bar{X}_{(\sigma,i)}\}_{\sigma \in \mathbb{N}^n}$, $i \in \mathbb{N}$, are independent and each $\{\bar{X}_{(\sigma,i)}\}$ is distributed as $\{X_\sigma\}$. We may assume $\{\bar{X}_{(\sigma,1)}\} = \{X_\sigma\}$. Then for each set S of subsets of $\mathbb{N} \times n$, if $S' = \{s \cup (\{n+1\} \times 1) : s \in S\}$, then

$$\mathcal{G}^{\bar{X}}_{S'} \subseteq \mathcal{G}^X_S \vee \mathcal{F}(\{\bar{X}_\tau : \tau(n+1) \neq 1\}).$$

But $\{X_\sigma\}$ is independent of $\{\bar{X}_\tau : \tau(n+1) \neq 1\}$ so

$$\mathcal{D}(\{X_\sigma\} | \mathcal{G}^{\bar{X}}_{S'}) = \mathcal{D}(\{X_\sigma\} | \mathcal{G}^X_S) \quad \text{a.s.}$$

Now (a) and (b) for $\{X_\sigma\}$ follow from (a) and (b) for $\{\bar{X}_\tau\}$. ■

This corollary generalizes Lemma 3.6 of Aldous [1]. The following corollary generalizes Aldous's Cor. 3.12.

Let us make the convention that for $t \subseteq n$, an array $\{X_\sigma\}_{\sigma \in \mathbb{N}^t}$ will be identified with the array $\{X'_\tau\}_{\tau \in \mathbb{N}^n}$ given by $X'_\tau = X_\sigma$ where σ is the restriction of τ to t.

5.11. Corollary: (a) Let U be a relation on Ω^n with support in $t \subsetneq n$. The following are equivalent

(1) The induced array $\{U_\sigma : \sigma \in \mathbb{N}^t\}$ on $\Omega^{\mathbb{N} \times t}$ is i.i.d.

(2) The induced array $\{U_\sigma\}$ on $\Omega^{\mathbb{N} \times n}$ is independent of $\mathscr{F}_S$, $S = \{s \cup \mathbb{N} \times (n-t) : s \subseteq \mathbb{N} \times t, \ |s| < |t|\}$.

(3) $\mathscr{G}(U)_S$ is trivial.

(b) If U^i $i \leq k$, are relations on Ω^n with support in $t_i \subseteq n$, such that $t_i \neq t_j, 1 \leq i < j \leq k$, and for each i, the array $\{U^i_\sigma\}$ on $\Omega^{\mathbb{N} \times t_i}$ is i.i.d., then $\{(U^1, \ldots, U^k)_\sigma\}_{\sigma \in \mathbb{N}^n}$ is independent of $\mathscr{F}_S$, where

$$S = \{s : |s \cap (1 \times t_i)^\pi| < |t_i|, i \leq k \text{ for every } \pi \text{ a finite, strict permutation of } \mathbb{N} \times n\}.$$

Proof: (a) (3) implies (2) by Corollary 5.9. (2) implies (1) by Corollary 5.6(b). (1) implies (3), because if σ_i, $i \in \mathbb{N}$ is any enumeration of $\mathbb{N}^t$, $\mathscr{G}_S$ is contained in the tail of $\{U_{\sigma_i} : i \in \mathbb{N}\}$.

(b) We will prove this by induction on k. We may assume that $t_k \not\subseteq t_j$ for any $j < k$. It follows that $|t_j \cap t_k| < |t_k|$ for each $j < k$. Hence if $T = \{s \cup \mathbb{N} \times (n-t_k) : |s| < |t_k|\}$, then $S \subseteq T$, and $t_j \in T$, $j < k$, so $\mathscr{F}_S \subseteq \mathscr{F}_T$ and $\{(U^1, \ldots, U^{k-1}_\sigma\}$ is $\mathscr{F}_T$-measureable. Hence, by (a), $\{U^k_\sigma\}$ is independent of $\mathscr{G}(U^1, \ldots, U^{k-1}) \vee \mathscr{F}_S$. By inductive hypothesis, $(U^1, \ldots, U^{k-1})$ is independent of $\mathscr{F}_S$. Therefore the result follows by the following lemma.

5.12. Lemma: Let X, Y be random variables. If X is independent of $\mathcal{F}(Y) \vee \mathcal{H}$, and Y is independent of $\mathcal{H}$, then (X, Y) is independent of $\mathcal{H}$.

The proof of this is obvious. ■

While we are on the subject of relations inducing i.i.d. arrays, let us give this useful result:

5.13. Proposition: Let $U : \Omega^{n+1} \to [0,1]$ be a relation on $(\Omega, \mathcal{F}_m, P_m)_{m \in \mathbb{N}}$ with U independent of $\mathcal{F}_{C(n+1,n)}$, and let $F_k \in \mathcal{F}_n$, $k \in \mathbb{N}$. Then for almost all ω, $U(\cdot, \omega) : \Omega^n \to [0,1]$ is independent of $\mathcal{F}_{C(n,n-1)}(F_k : k \in \mathbb{N})$, and $P_n(\varphi(U(\cdot, \omega))) = P_{n+1}(\varphi(U))$ for every $\varphi \in C([0,1])$.

Proof: It will suffice to prove this where there are only finitely many F_k, $k \leq M$. But in this case we may take $\{F_k\}$ to be a partition of Ω^n. Then if $G \in \mathcal{F}_{C(n,n-1)}$, $F \in \mathcal{F}(F_k : k \leq M)$, $\varphi \in C([0,1])$, then

$$P_n(\varphi(U(\cdot, \omega))GF) = \sum_{F_k \subseteq F} P_n(\varphi(U(\cdot, \omega))GF_k) \qquad \text{a.s.}(\omega)$$

so it will suffice to prove for a single F that for each G, φ,

$$P_n(\varphi(U(\cdot, \omega))GF\ ) = P_n(\varphi(U(\cdot, \omega))P_n(GF\ ) \qquad \text{a.s.}(\omega)$$

If this fails we may take it that $P(B) > 0$, where

$$B = \{\omega : P_n(\varphi(U(\cdot, \omega))GF) > P_n(\varphi(U(\cdot, \omega)))P_n(GF)\}\ .$$

But $B \times GF \in \mathcal{F}_{C(n+1,n)}$, and

$$P_{n+1}(\varphi(U)(B\times(GF))) - P_{n+1}(\varphi(U))P_{n+1}(B\times(GF))$$

$$= \int_B (P_n(\varphi(U(\cdot,\omega))GF) - P_n(\varphi(U(\cdot,\omega))P_n(GF))dP_1(\omega)$$

$$> 0$$

But this contradicts the hypothesis that U is independent of $\mathcal{F}_{C(n+1,\, n)}$.
What I have shown of course includes the fact that for any $\varphi \in C([0,1])$
$P_n(\varphi(U(\cdot,\omega))) = P_{n+1}(\varphi(U))$ for almost all ω, hence for all $\varphi \in C([0,1])$,
since we need only consider φ in a countable dense set. ∎

6. Atomless Spaces.

We are now in a position to directly give a proof of the structural characterization of ED and DRE arrays which we are after, but we will not do this. Instead, we shall first introduce some measure-algebraic material which will enable us to prove a stronger theorem with a slightly simpler proof; in any case, it seems necessary to introduce some measure theory in order to characterize equivalent representations.

The results summarized in this section are adaptations of well-known facts about atomless measure spaces. These are fairly simple, and must be folklore, but we do not know a source for them.

6.1. Definition: Let $(\Omega, \mathcal{F}, P)$ be a probability space $\mathcal{G}$ a σ-algebra contained in $\mathcal{F}$.

(a) $\mathcal{F}$ is said to be atomless over $\mathcal{G}$ if for every $F_1, \ldots, F_n \in \mathcal{F}$ and $\mathcal{G}(F_1, \ldots, F_n)$-measureable function $f: \Omega \to [0,1]$, there is $F \in \mathcal{F}$ such that

$$P(F \mid \mathcal{G}(F_1, \ldots, F_n)) = f \qquad \text{a.s.}$$

(b) Let κ be an infinite cardinal. $\mathcal{F}$ is κ-atomless over $\mathcal{G}$ iff whenever $F_i \in \mathcal{F}$, $i \in I$, with I of cardinality less than κ, $\mathcal{F}$ is atomless over $\mathcal{G}(F_i : i \in I)$. (Observe that atomless = $\aleph_0$-atomless, where $\aleph_0$ is the cardinal of $\mathbb{N}$.)

(c) Let us say that $U_i \in \mathcal{F}$, $i \in I$ are independent over $\mathcal{G}$ if $\mathcal{G}$, U_i, $i \in I$ are mutually independent. ■

6.2. Terminology: A well-known theorem of von Neumann (see Royden [13]) says that whenever M_1, M_2 are complete separable metric spaces and μ_1, μ_2 are atomless measures on the Borel sets of M_1, M_2, respectively, then there is a bijection $\varphi : M_1 \to M_2$ such that φ, φ^{-1} are Borel measureable, and measure preserving with respect to the given measures. Thus if X is an M_1-valued random variable there is an M_2-valued random variable $Y = \varphi(X)$ which is measure-algebraically equivalent to X, and if the distribution of X is atomless, we may choose Y to have any atomless distribution on M_2. We shall refer to the process of obtaining Y from X as recoding. ∎

The various degrees of "atomless over" have this simple characterization:

6.3. Proposition: Let $\mathcal{N}$ denote the nullsets of P. $\mathcal{F}$ is atomless over $\mathcal{G}$ iff for every $F \in \mathcal{F}$, $F \notin \mathcal{N}$, the principal ideal $(F) = \{H \in \mathcal{F} : H \subseteq F\}$ is not generated (as a σ-algebra) by less than κ elements over (i.e. together with) $(F) \cap (\mathcal{G} \vee \mathcal{N})$.

Proof: (⟹) is trivial. (⟸) follows by a simple modification of the proof of Maharam [12], Lemma 2. ∎

The following characterization in terms of sets independent over $\mathcal{G}$ will be the most useful in connection with g.p.s.'s.

6.4. Proposition: $\mathcal{F}$ is atomless over $\mathcal{G}$ iff there are U_n, $P(U_n) = r \in (0,1)$, $n \in \mathbb{N}$, independent over $\mathcal{G}$.

Proof: (⟹) follows easily by iterated application of the defining property of "atomless over".

($\Leftarrow$ Suppose to the contrary that for some $F \in \mathcal{F}$, $P(F) > 0$, (F) is generated by $H_1, \ldots, H_n$ over $\mathcal{G} \cap (F) \vee \mathcal{H}$. We may as well assume that $H_1, \ldots, H_m$ form a partition of F and $P(H_1) > 0$. Thus there are $G_n \in \mathcal{G}$ such that

Thus
(1) $H_1 U_n = H_1 G_n$ a.s., $n \in \mathbb{N}$.

(2) $H_1 U_n = H_1 U_n G_n$ a.s., $n \in \mathbb{N}$.

But by the weak law of large numbers,

$$(*) \quad \lim_n P\left(\frac{1}{N} \sum_{n \leq N} H_1 U_n\right) = r P(H_1)$$

$$= \lim_n P\left(\frac{1}{N} \sum_{n \leq N} H_1 U_n G_n\right) \quad \text{by (2)}$$

$$= r \lim_n \frac{1}{N} \sum_{n \leq N} P(H_1 G_n)$$

by WLLN conditional on $\mathcal{G}$. But

$$(*) = \lim_n \frac{1}{N} \sum_{n \leq N} P(H_1 G_n), \quad \text{by (1)}.$$

Thus $P(H_1) = r P(H_1)$, a contradiction, since $0 < r < 1$. ∎

Now we work in the converse direction, to show that when $\mathcal{F}$ is atomless over $\mathcal{G}$, any $\mathcal{F}$-measureable function can be written in terms of $\mathcal{G}$ and sets independent over $\mathcal{G}$.

6.5. Lemma. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{F}$ atomless over $\mathcal{G} \subseteq \mathcal{F}$. Then for any $F \in \mathcal{F}$, $\varepsilon > 0$, there are $U_1, \ldots, U_m \in \mathcal{F}$ independent over $\mathcal{G}$, such that for some $Y \in \mathcal{G}(U_i : i < m)$, $P(|F - Y|) < \varepsilon$.

Proof: For $r \in [0,1]$ $m \in \mathbb{N}$, let $[r]^m$ denote the greatest number $\frac{k}{m}$, $k \in \mathbb{N}$, which is no larger than r. Given ε, let m be the least integer such that $P(|g-[g]^m|) < \varepsilon$, where $g = P(F|\mathcal{G})$. Since $\mathcal{F}$ is atomless over $\mathcal{G}$, we may choose in turn U_k, $1 \leq k < m$, so that

$$\begin{aligned} P(U_k|\mathcal{G}(F, U_j : j < k)) &= \frac{k}{mg} \quad \text{on } F \cap \{g \in [\tfrac{k}{m}, \tfrac{k+1}{m})\} \\ &= \frac{k}{m} \quad g \notin [\tfrac{k}{m}, \tfrac{k+1}{m}) \\ &= 0 \quad \text{elsewhere.} \end{aligned}$$

Then

$$P(U_k|\mathcal{G}(U_j:j<k)) = \frac{k}{mg}\{g \in [\tfrac{k}{m}, \tfrac{k+1}{m})\}P(F|\mathcal{G}(U_j:j<k)) + \frac{k}{m}\{g \notin [\tfrac{k}{m}, \tfrac{k+1}{m})\} .$$

But it follows from the definition of U_j, $j < k$, that each $V_j = \{g \in [\frac{k}{m}, \frac{k+1}{m})\}U_j$ is c.i. of $F, U_i, i < j$, given $\mathcal{G}$, and hence V_j, $j < k$, and F are c.i. given $\mathcal{G}$. But

$$\{g \in [\tfrac{k}{m}, \tfrac{k+1}{m})\}P(F|\mathcal{G}(V_j:j<k)) = \{g \in [\tfrac{k}{m}, \tfrac{k+1}{m})\}P(F|\mathcal{G}(U_j:j<k))$$

hence

$$\begin{aligned} P(U_k|\mathcal{G}(U_j:j<k)) &= \frac{k}{mg}\{g \in [\tfrac{k}{m}, \tfrac{k+1}{m})\}g + \frac{k}{m}\{g \notin [\tfrac{k}{m}, \tfrac{k+1}{m})\} \\ &= \frac{k}{m} . \end{aligned}$$

Thus U_k is independent of $\mathcal{G}(U_j:j<k)$ for each $k < m$, and it follows that U_k, $k < m$, are independent over $\mathcal{G}$.

Now if

$$Y = U_k \quad \text{if} \quad g \in [\frac{k}{m}, \frac{k+1}{m}) \quad 1 \le k \le m,$$
$$= 1 \quad \text{if} \quad g = 1$$
$$= 0 \quad \text{if} \quad g = 0$$

then $Y \in \mathcal{G}(U_k : k<m)$, $Y \le F$, and

$$P(|F-Y|) = P(P(F-Y|\mathcal{G}))$$
$$= P(g-[g]^k) < \varepsilon \ ,$$

as was desired. ∎

6.6. Definition: A random variable is uniform if it is uniformly distributed on $[0,1]$. ∎

6.7. Proposition: Any $\mathcal{F}$-measureable random variable is of the form $X = f(Z, U)$, where Z is $\mathcal{G}$-measureable, and U is uniform and independent of $\mathcal{G}$.

Proof: Recode X as $\overline{X} : \Omega \to \{0,1\}^{\mathbb{N}}$, $\overline{X} = (\overline{X}_i : i \in \mathbb{N})$. Enumerate $\mathbb{N}$ as i_n $n \in \mathbb{N}$ so that each $i \in \mathbb{N}$ appears as i_n for infinitely many n. Now proceed inductively: Suppose that before the n^{th} step we have chosen m_j, $j < n$, U_k, $k \le m_{n-1}$, independent over $\mathcal{G}$, such that $P(U_{m_j}) = \frac{1}{2}$, $j < n$, and $P(|\overline{X}_{ij} - Y_j|) < \frac{1}{j}$ for some $Y_j \in \mathcal{G}(U_1, \ldots, U_{m_j})$, $j \le n-1$. Now choose $m_n > m_{n-1}$, U_k, $m_{n-1} < k < m_n$ independent over $\mathcal{G}$ such that

$$P(|\overline{X}_{i_n} - Y_n|) < \frac{1}{n}, \quad \text{for some } Y_n \in \mathcal{G}(U_1, \ldots, U_{m_{n-1}}),$$

and choose U_{m_n} independent of $\mathscr{A}(U_1, \ldots, U_{m_n - 1})$, $P(U_{m_n}) = \frac{1}{2}$. But $U_1, \ldots, U_{m_n}$ are independent over $\mathscr{A}$. Thus we construct U_n, $n \in \mathbb{N}$, such that for every $i \in \mathbb{N}$, $\overline{X}_i$ is approximated arbitrarily closely by sets in $\mathscr{A}(U_n : n \in \mathbb{N})$, and for each n, $P(U_{m_n}) = \frac{1}{2}$. Hence there is $\overline{Y}_i \in \mathscr{A}(U_n : n \in \mathbb{N})$ such that $\overline{X}_i = \overline{Y}_i$ a.s. Recoding $(\overline{Y}_i : i \in \mathbb{N})$ we get Y such that $X = Y$ a.s. But $\overline{U} = (U_i : i \in \mathbb{N})$ has an atomless distribution on $\{0,1\}^{\mathbb{N}}$ (since $P(U_i) = \frac{1}{2}$ for infinitely many i), hence may be recoded as a uniform random variable U. But then $Y = f(Z, U)$ for some Borel measureable f, $\mathscr{A}$-measureable Z. ∎

7. Structure Theory.

We now apply the results of the last two chapters to g.p.s.'s and relations on them to obtain a structural characterization of the relations in terms of relations independent of sets with support smaller than their own. This yields by sampling a representation theorem for ED and DRE arrays in terms of i.i.d. uniform arrays; this can be extended to RCE and RE arrays. We also determine when two different representations produce equivalent arrays.

Let $(\Omega, \mathcal{F}_n, P_n)_{n\in\mathbb{N}}$ be a g.p.s. and let us write $\mathcal{F}_{<n}$ for $\mathcal{F}_{C(n,n-1)}$ $n \in \mathbb{N}$, and analogously $\mathcal{F}_{<t}$ for $\mathcal{F}_{C(t,t-1)}, t \subseteq m \in \mathbb{N}$ ($\mathcal{F}_{<1} = \{\phi, \Omega\}$). We want to work in the situation where $\mathcal{F}_m$ is atomless or $\aleph_1$-atomless over $\mathcal{F}_{<m}$ (where $\aleph_1$ is the least uncountable cardinal--i.e. $\mathcal{F}_m$ is atomless over $\mathcal{F}_{<m}(F_k : k\in\mathbb{N})$ whenever $F_k \in \mathcal{F}_m, k \in \mathbb{N}$) for each $m \leq n$. The following results shows that it is easy to arrange this.

7.1. Proposition: Let $(\Omega, \mathcal{F}_n, P_n)_{n\in\mathbb{N}}$ be a g.p.s. If there is $U \in \mathcal{F}_n$, $1 > P(U) > 0$, which is independent of $\mathcal{F}_{<n}$ (in particular, if $\mathcal{F}_n$ is atomless over $\mathcal{F}_{<n}$), then $\mathcal{F}_m$ is $\aleph_1$-atomless over $\mathcal{F}_{<m}$, for each $m < n$.

Proof: Suppose we are given $F_k \in \mathcal{F}_m$, $k \in \mathbb{N}$. Then by Prop. 5.13 we may successively choose $\vec{\omega}_i \in \Omega^{n-m}, i \in \mathbb{N}$, such that $U_i = U(\cdot, \vec{\omega}_i)$ has $P_n(U) = P_m(U(\cdot, \vec{\omega}_i))$ and U_i is independent of $\mathcal{F}_{<m}(F_k, U_j : k \in \mathbb{N}, j < i)$. Then U_i, $i \in \mathbb{N}$, are independent over $\mathcal{F}_{<n}(F_k : k\in\mathbb{N})$, so the result follows by 6.4. ■

7.2. Definition: (a) Let $U^t, t \subseteq n$ be a family of uniform relations on Ω^n (i.e. uniformly distributed random variables with values in $[0,1]$). We

We call $\{U^t : t \subseteq n\}$ an independent family if for each t, U^t has support in t and $\{(U^t)^{ED}_\sigma\}_{\sigma \in \mathbb{N}^t}$ is i.i.d. It follows easily from Cor. 5.11(b) and Cor. 5.6(b) that $\{(U^t)^{ED}_\sigma\}_{\sigma \in \mathbb{N}^n}$, $t \subseteq n$, are mutually independent.

(b) If also $(U^t)^\pi = U^{\pi(t)}$ for each $\pi \in n!$, $t \subseteq n$ then we say $\{U^t : t \subseteq n\}$ is symmetric independent. It follows that $\{(U^t)^{DWE}_s\}_{s \in C(\mathbb{N}, t)}$ $t \subseteq n$, are each i.i.d., and independent of each other. This of course determines the distribution of $\{(U^t : t \subseteq n)^{DRE}_\sigma\}_{\sigma \in \mathbb{N}^n}$.

(c) If $\{\{V^t_\sigma\} : t \subseteq n\}_{\sigma \in \mathbb{N}^n}$ are ED (resp. DRE) and have the same distribution as $\{(U^t : t \subseteq n)^{ED}_\sigma\}$ (resp. $\{(U^t : t \subseteq n)^{DRE}_\sigma\}$), where $\{U^t : t \subseteq n\}$ is an independent family (resp. symmetric independent) then we will call $\{\{V^t_\sigma\} : t \subseteq n\}$ an independent (resp. symmetric independent) family of arrays.

(d) If $\{V^t : t \subseteq n\}$ is also an independent (or symmetric independent) family of relations which are independent of $\{U^t : t \subseteq n\}$ we will say that $\{U^t, V^t : t \subseteq n\}$ is jointly (symmetric) independent. $\{(U^t, V^t) : t \subseteq n\}$ actually satisfies the conditions of the definition of (symmetric) independence, except that each random variable (U^t, V^t) is uniformly distributed on $[0,1]^2$. ∎

Since any two symmetric independent families induce DRE arrays with the same joint distributions, writing a relation as

$$X = f(U^t : \phi \neq t \subseteq n)$$

determines the distribution of $\{X^{DRE}_\sigma\}$.

7.3. Theorem: (a) If $\mathcal{F}_n$ is atomless over $\mathcal{F}_{<n}$, then for any relation X on Ω^n,

$$X = f(U^t : \phi \neq t \subseteq n) \quad \text{a.s.}$$

where f is Borel measureable and $\{U^t : t \subseteq n\}$ is a symmetric independent family.

(b) If $\mathcal{F}_n$ is $\aleph_1$-atomless over $\mathcal{F}_{<n}$ and $\{U^t : t \subseteq n\}$ is independent (or symmetric independent) then for any relation X on Ω^n, there are $V^t, t \subseteq n$ such that $\{U^t, V^t : t \subseteq n\}$ is jointly independent (or symmetric independent) and

$$X = f(U^t, V^t : \phi \neq t \subseteq n)$$

for some Borel measureable f.

Proof: Let us start with the non-symmetric case of (b). We will prove by induction on $m \leq n$, that given relations Z^s, $s \in C(n, m)$, each with support in s, there are V^t, $t \subseteq n$, $|t| \leq m$, such that $\{U^t, V^t : t \subseteq n, |t| \leq m\}$ is jointly independent (except that it is not defined for all $t \subseteq n$) and g^s, $s \in C(n, m)$, Borel measureable such that $Z^s = g^s(U^t, V^t : \phi \neq t \subseteq s)$ for each s. (Thus the theorem is the case $m = n$, $X = Z^n$, $f = g^n$). By Prop. 7.1, $\mathcal{F}_m$ is $\aleph_1$-atomless over $\mathcal{F}_{<m}$, and so is $\mathcal{F}_s$ over $\mathcal{F}_{<s}$, $s \in C(n, m)$ since $(\mathcal{F}_m, \mathcal{F}_{<m})$ is isomorphic to $(\mathcal{F}_s, \mathcal{F}_{<s})$. Thus by Prop. 6.7 we can write $Z^s = h^s(\tilde{W}^s, V^s)$ where $\tilde{W}^s$ is $\mathcal{F}_{<s}(U^s)$ measureable, and V^s is uniform and independent of $\mathcal{F}_{<s}(U^s)$. But for each s,

$$\tilde{W}^s = k^s (W^{su}, U^s : u \subseteq s, |u| = m - 1),$$

where W^{su} have support in u. For each u, we can recode $(W^{su} : s \supseteq u)$ into a single $[0,1]$-valued random variable W^u, so we may as well assume

$$\tilde{W}^s = k^s\,(W^u, U^s : u \subseteq s,\ |u| = m-1)\ .$$

But by induction each

$$W^u = g^u(U^t, V^t : \phi \neq t \subseteq u)$$

for some partial jointly independent $\{U^t, V^t : |t| \leq m-1\}$. But by Cor. 5.11(b) $\{U^t, V^t : t \subseteq m\}$ is jointly independent, and for each s,

$$Z^s = h^s(k^s(g^u(U^t, V^t : \phi \neq t \subseteq u),\ U^s : u \subsetneq s), V^s)$$

$$= g^s(U^t, V^t : \phi \neq t \subseteq s)\ .$$

as desired.

The proof of the symmetric case of (b) needs an extra wrinkle to get $\{V^t : t \subseteq n\}$ to be symmetric, but because $\{U^t : t \subseteq n\}$ is symmetric, a simpler induction will do. We show by induction on n that the statement is true of any relation X. By Prop. 7.1, $\mathcal{F}_1$ is atomless, hence there is a uniform $\mathcal{F}$-measureable $Y : \Omega \to [0,1]$. Now if for each $\pi \in n!$,

$$Z(\pi)(\omega_1, \ldots, \omega_n) = X(\omega_{\pi^{-1}(1)}, \ldots, \omega_{\pi^{-1}(n)}) \text{ if } Y(\omega_1) < \ldots < Y(\omega_n)$$

$$= 0 \qquad \text{otherwise,}$$

then $X = \sum_{\pi \in n!} Z(\pi)^{\pi}$ a.s. Hence it will suffice to prove the result for $Z = (Z(\pi) : \pi \in N!)$. By recoding and Prop. 6.7, we can write

$$Z = g(W^i, U^n, \tilde{V} : i \leq n) \text{ a.s.}$$

where W^i has support in $n - \{i\}$, and $\tilde{V}$ is uniform and independent of $\mathcal{F}_{<n}(U^n)$. By recoding, we may suppose $Y^i(\vec{\omega}) = Y(\omega_i)$ is a function of W^{i+1} (W^1, if $i = n$) for each i. Since Z is zero unless $Y(\omega_1) < \ldots < Y(\omega_n)$, if we let

$$V^n(\omega_1, \ldots, \omega_n) = \tilde{V}(\omega_{\pi(1)}, \ldots, \omega_{\pi(n)}) \text{ if there is a (necessarily unique) } \pi \in n! \text{ such that } Y(\omega_{\pi(1)} < \ldots < Y(\omega_{\pi(n)})$$

$$= 0 \qquad \text{otherwise,}$$

then

$$Z = h(W^i, U^n, V^n : i < n) \quad \text{a.s.}$$
$$= \begin{cases} g(W^i, U^n, V^n : i \leq n) & \text{if } Y^1 < \ldots < Y^n \\ 0 & \text{otherwise .} \end{cases}$$

Now V^n is uniform and symmetric, and since $\mathcal{F}_{<n}(U^n)^\pi = \mathcal{F}_{<n}((U^n)^\pi) = \mathcal{F}_{<n}(U^n)$ for each $\pi \in n!$, it is also independent of the latter σ-algebra. Hence (U^n, V^n) is symmetric, uniform on $[0, 1]^2$, and independent of $\mathcal{F}_{<n}$. For $i \leq n$, let $\pi_i \in n!$ be the permutation which transposes i and n and fixes everything else, and let $\overline{W}^i = (W^i)^{\pi_i}$.

By induction and recoding, there is $\{V^t : t \subseteq n - 1\}$ such that $\{U^t, V^t : t \subseteq n - 1\}$ is jointly symmetric independent, and

$$\overline{W}^i = f^i(U^t, V^t : \phi \neq t \subseteq n - 1) \qquad i \leq n .$$

If for $t \subseteq n$, $|t| < n$, we define $V^t = (V^{\pi(t)})^{\pi^{-1}}$ for any π such that $\pi(t) \subseteq n - 1$, then, by symmetry of $\{V^t : t \subseteq n - 1\}$, the new V^t's are well-defined, and $\{U^t, V^t : t \subseteq n\}$ is symmetric independent. Thus

$$Z = h(f^i(U^{\pi_i^{-1}(t)}, V^{\pi_i^{-1}(t)} : t \subseteq n - 1), \ U^n, V^n : i \leq n)$$

$$= f(U^t, V^t : \phi \nsubseteq t \subseteq n)$$

Part (a) is similar, but a little simpler. ■

7.4. Corollary: (a) $\{X_\sigma\}_{\sigma \in \mathbb{N}^n}$ is ED iff there is a independent family of arrays $\{U^t_\sigma : t \subseteq n\}$ such that $\{X_\sigma\} \sim \{Y_\sigma\}$ where

$$Y_\sigma = f(U^t_\sigma : \phi \nsubseteq t \subseteq n) ,$$

f Borel measureable.

(b) $\{X_\sigma\}$ is DRE iff $\{X_\sigma\} \sim \{Y_\sigma\}$ for

$$Y_\sigma = f(U^t_\sigma : \phi \nsubseteq t \subseteq n),$$

f Borel measureable, $\{(U^t_\sigma : t \subseteq n)\}$ symmetric independent

(c) $\{X_s\}_{s \in C(\mathbb{N}, n)}$ is DWE iff $\{X_s\} \sim \{Y_s\}$

$$Y_s = f(U^t_s : \phi \nsubseteq s \subseteq n)$$

$\{(U^t_s : t \leq n)\}$ symmetric independent, and

$$f(x^{\pi(t)} : \phi \nsubseteq t \subseteq n) = f(x^t : \phi \nsubseteq t \subseteq n)$$

for each $\pi \in n!$,

Proof: (<==) is trivial in all three cases. (==>)

(a) Let $\{\bar{U}_\sigma\}_{\sigma \in \mathbb{N}^n}$ be i.i.d. and independent of $\{X_\sigma\}$, with each $\bar{U}_\sigma$ uniform. Then $\{(X_\sigma, \bar{U}_\sigma)\}$ is ED. By the Sampling Representation, $\{(X_\sigma, \bar{U}_\sigma)\} \sim \{(Y_\sigma, U_\sigma)\}$, where $\{(Y_\sigma, U_\sigma)\}$ is obtained by ED sampling of relations Y, U on Ω^n for some g.p.s. $(\Omega, \mathcal{F}_n, P_n)$. By Cor. 5.4(a), U is independent of $\mathcal{F}_{<n}$. Recoding U as a uniformly distributed $\{0,1\}^{\mathbb{N}}$-valued relation $(U_i : i \in \mathbb{N})$, $U_i \in \mathcal{F}_{<n}$ are independent of $\mathcal{F}_{<n}$, and $P_n(U_i) = \frac{1}{2}$. Hence by Prop. 6.4, $\mathcal{F}_{<n}$ is atomless over $\mathcal{F}_{<n}$. Hence by 7.3 $Y = f(U^t : \phi \neq t \subseteq n)$ as for some independent family $\{U^t : t \subseteq n\}$ and f Borel measureable. If $\{U^t_\sigma\} = \{(U^t)^{ED}_\sigma\}$ for each $t \subseteq n$, then $\{(U^t_\sigma : t \subseteq n)\}$ is independent, and

$$Y_\sigma = f(U^t_\sigma : \phi \neq t \subseteq n) \text{ a.s.} \qquad \sigma \in \mathbb{N}^n,$$

and the result follows by changing Y_σ on a set of measure zero.

(b) Is similar.

(c) By (b), we may assume that $\{X_\sigma\}$, the DRE array induced by $\{X_s\}$, is induced by some relation $X = g(U^t : t \subseteq n)$ where $\{U^t, t \subseteq n\}$ is, symmetric independent.

Then $X^\pi = X$ a.s. for each $\pi \in n!$, so if

$$f(x^t : t \subseteq n,\ t \neq \phi) = \frac{1}{n!} \sum_{\pi \in n!} g(x^{\pi(t)} : t \subseteq n,\ t \neq \phi),$$

then

$$Y_s = f(U_s^t : \phi \neq t \subseteq n) = X_s \quad \text{a.s.}$$

for each s, and f has the desired symmetries. ■

7.5. Corollary: $\{X_\sigma\}$ is RCE (RE) iff $\{X_\sigma\} \sim \{Y_\sigma\}$ where

$$Y_\sigma = f(U_\sigma^t : t \subseteq n) \qquad \text{(i.e. including } t = \phi)$$

where f is Borel,
$\{(U_\sigma^t : \phi \neq t \subseteq n)\}$ is independent (symmetric independent) and $U_\sigma^\phi = U^\phi$ is uniform and independent of $\{(U_\sigma^t : t \subseteq n)\}$. The analogous result for WE arrays also holds.

Proof: We use the same device as in Cor. 5.10(c). Let us assume $\{X_\sigma\}_{\sigma \in \mathbb{N}^n}$ is RCE; RE is similar and WE follows from RE as DWE from DRE in the theorem. Form $\{\overline{X}_\tau\}_{\tau \in \mathbb{N}^{n+1}}$ as in Cor. 5.10(c). Since it is ED, $\{\overline{X}_\tau\} \sim \{\overline{Y}_\tau\}$ where

$$\overline{Y}_\tau = \overline{f}(\overline{U}_\tau^t : \phi \neq t \subseteq n+1)$$

as in the theorem. But $\{X_\sigma\} \sim \{Y_\sigma\} = \{\overline{Y}_{(\sigma,1)}\}_{\sigma \in \mathbb{N}^n}$, and

$$Y_\sigma = \overline{f}(\overline{U}_{(\sigma,1)}^t : \phi \neq t \subseteq n+1).$$

Now $\overline{U}_{(\sigma,1)}^{\{n+1\}}$, $(U_{(\sigma,1)}^t, U_{(\sigma,1)}^{t \cup \{n+1\}})$, $\phi \neq t \subseteq n, \sigma \in \mathbb{N}^n$ satisfy what the Corollary requires of U_σ^t, $t \subseteq n$ $\sigma \in \mathbb{N}^n$, except that $(\overline{U}_{(\sigma,1)}^{\{n+1\}}, \overline{U}_{(\sigma,1)}^{t \cup \{n+1\}})$, $t \subseteq n$, $\sigma \in \mathbb{N}^n$ are uniformly distributed on $[0,1]^2$. We can recode them as $\{(U_\sigma^t : \phi \neq t \subseteq n)\}$

(independent) and let $U^{\phi} = \overline{U}_{(\sigma,1)}^{\{n+1\}}$. Then

$$Y_\sigma = f(U_\sigma^t : t \subsetneq n), \qquad \sigma \in \mathbb{N}^n,$$

as desired. ■

7.6. Corollary: If $(\Omega, \mathcal{F}_m, P_m)_{m\in\mathbb{N}}$ is a g.p.s. with $\mathcal{F}_n$ atomless over $\mathcal{F}_{<n}$, and $\{X_\sigma\}_{\sigma\in P(\mathbb{N},n)}$ is DRE (or $\{X_\sigma\}_{\sigma\in\mathbb{N}^n}$ is ED) then there is a relation Y on Ω^n with $\{X_\sigma\} \sim \{Y_\sigma^{DRE}\}$ $(\{X_\sigma\} \sim \{Y_\sigma^{ED}\})$.

Proof: By Cor. 7.5 we can assume that for each σ, $X_\sigma = f(U_\sigma^t : \phi \neq t \subseteq n)$, where $\{(U_\sigma^t : t \subseteq n)\}$ is symmetric independent. It is implicit in Theorem 7.3(a) that there is a symmetric independent family $\{V^t : t \subseteq n\}$ on Ω^n. Put $Y = f(V^t : t \subseteq n)$ (ED case similar). ■

Corollaries 7.4 and 7.5 generalize Theorem 1.4 and Prop. 3.3 of Aldous [1]. In any of the cases covered by those corollaries, we will call $f(U_\sigma^t : t \subseteq n)$ a (structural) representation of $\{X_\sigma\}$.

Two natural questions about structural representations arise.

(1) When does a relation X have a representation $f(U^t : t \in S)$ for some given set S of subsets of n?

(2) When do two representations induce arrays with the same distribution? (Asked by Aldous)

We can give a satisfactory answer to (2), but we do not have a good answer to (1); perhaps some light would be shed on the matter by careful consideration of the answer to (2). But we can make some simple observations about (1).

The first is that if S is a set such that $t \subseteq s \in S$ implies $t \in S$, then if X is $\mathscr{A}(X)_S$ measureable, then clearly X has a representation $f(U^t : t \in S)$. One might hope that a sort of inversion of this, that if X is independent of $\mathscr{A}(X)_S$, then X has a representation $f(U^t : t \notin S)$, would hold, but it does not. Any ideas in this line are foiled by counterexamples of the following simple kind (elaborating one used by Aldous): Let

$$X = f(U^1, U^{23}) = 1 \text{ if } U^1, U^{23} \geq \frac{1}{2} \text{ or } U^1, U^{23} \leq \frac{1}{2}$$

$$= 0 \qquad \text{otherwise}$$

Then

$$P(X|\mathscr{F}_1) = \int_2^1 f(U^1, x)dx = \frac{1}{2} = P(X)$$

so X is independent of $\mathscr{F}_1$; it is likewise independent of $\mathscr{F}_{\{2,3\}}$. But as it is $\mathscr{F}_{\{\{1\},\{2,3\}\}}$-measureable, it must be represented in some form like that it is in.

Furthermore X may have a representation in which one variable does not appear, and another in which it occurs nontrivially. Thus if f is as before, and

$$g(x) = 1 \text{ if } x \geq \frac{1}{2}$$

$$= 0 \text{ if } x < \frac{1}{2} .$$

then $f(U^1, U^{12})$ and $g(U^{12})$ induce ED arrays with the same distribution.

Before answering the second question we will prove a further fact about g.p.s.'s.

7.7. Theorem. Let $(\Omega, \mathcal{F}_m, P_m)_{m\in\mathbb{N}}$ be a g.p.s. such that $\mathcal{F}_n$ is $\aleph_1$-atomless over $\mathcal{F}_{<n}$. If X is a relation on Ω^n, and $\{(\bar{X}_\sigma, \bar{Y}_\sigma)\}$ is a DRE array with $\{X_\sigma^{DRE}\} \sim \{\bar{X}_\sigma\}$ (resp. ED array with $\{X_\sigma^{ED}\} \sim \{\bar{X}_\sigma\}$) then there is a relation Y on Ω^n with $\{(X_\sigma^{DRE}, Y_\sigma^{DRE})\} \sim \{(\bar{X}_\sigma, \bar{Y}_\sigma)\}$ (resp. $\{(X_\sigma^{ED}, Y_\sigma^{ED})\} \sim \{(X_\sigma, Y_\sigma)\}$).

To prove this, we will need the following Lemma (which was proved in [7], though not explicitly stated).

7.8. Lemma: If $\{(X_\sigma^{(1)}, Y_\sigma^{(1)})\}_{\sigma\in\mathbb{N}^n}$ and $\{(X_\sigma^{(2)}, Z_\sigma^{(2)})\}_{\sigma\in\mathbb{N}^n}$ are DRE (resp. ED) arrays such that $\{X_\sigma^{(1)}\} \sim \{X_\sigma^{(2)}\}$ then there is a DRE (resp. ED) array $\{(X_\sigma, Y_\sigma, Z_\sigma)\}_{\sigma\in\mathbb{N}^n}$ such that $\{(X_\sigma, Y_\sigma)\} \sim \{(X_\sigma^{(1)}, Y_\sigma^{(1)})\}$ and $\{(X_\sigma, Z_\sigma)\} \sim \{(X_\sigma^{(2)}, Z_\sigma^{(2)})\}$.

Proof: Define the distribution P of an array $\{(\bar{X}_\sigma, \bar{Y}_\sigma, \bar{Z}_\sigma)\}_{\sigma\in\mathbb{N}^n} = (\bar{X}, \bar{Y}, \bar{Z})$ by

(1) $(\bar{X}, \bar{Y}) \sim (\bar{X}^{(1)}, \bar{Y}^{(1)})$

(2) $(\bar{X}, \bar{Z}) \sim (\bar{X}^{(2)}, \bar{Z}^{(2)})$

(3) $P(\bar{X} \in B_1, \bar{Y} \in B_2, \bar{Z} \in B_3)$

$$= P(\{\bar{X} \in B_1\} P(\bar{Y} \in B_2 | \bar{X}) P(\bar{Z} \in B_3 | \bar{X})).$$

B_i Borel sets in $\mathbb{R}^{\mathbb{N}^n}$ (i.e. combine $(X^{(1)}, Y^{(1)})$ and $(X^{(2)}, Z^{(2)})$ independently over the identification of $X^{(1)}$ and $X^{(2)}$). This defines a consistent distribution, which is clearly RE (resp. RCE). By the deFinetti theorem 2.2, P is a mixture of DRE (resp. ED) distributions Q; by the uniqueness of the mixture, since $(\bar{X}, \bar{Y})$ and $(\bar{X}, \bar{Z})$ are DRE, Q is the same as P on $(\overline{X}, \overline{Y})$ and $(\overline{X}, \overline{Z})$ almost surely with respect to the mixing measure. Choose such a Q as the distribution for a DRE (ED) array $\{(X_\sigma, Y_\sigma, Z_\sigma)\}_{\sigma \in \mathbb{N}^n}$. This array satisfies the conclusions of the Lemma. ■

Proof of 7.7: Let $X = f(U^t : \phi \neq t \subseteq n)$ a.s. where $\{U^t : t \subseteq n\}$ is symmetric independent. Amalgamate $\{(X, U^t : t \subseteq n)_\sigma^{DRE}\}$ and $\{(\overline{X}_\sigma, \overline{Y}_\sigma)\}$ as in the Lemma to get $\{(\tilde{X}_\sigma, \tilde{U}_\sigma^t, \tilde{Y}_\sigma : \phi \neq t \subseteq n)\}$ with the indicated marginal distributions: we may assume that this is the DRE array induced by relations $\tilde{X}, \tilde{U}^t, \tilde{Y}$, $\phi \neq t \subseteq n$, on some g.p.s. $(\tilde{\Omega}, \mathcal{A}_m, Q_m)$ with $\mathcal{A}_n$ $\aleph_1$-atomless over $\mathcal{A}_{<n}$. Thus we can write $\tilde{Y} = g(\tilde{U}^t \tilde{V}^t : \phi \neq t \subseteq n)$ a.s. where $\{\tilde{U}^t, \tilde{V}^t \ t \subseteq n\}$ is symmetric independent. But by $\aleph_1$-atomlessness there exist V^t on Ω^n satisfying the same properties as $\tilde{V}^t$. Then if we define $Y = g(U^t, V^t : \phi \neq t \subseteq n)$, $(X, Y)^{DRE} \sim (\overline{X}, \overline{Y})^{DRE}$. The ED case is similar. ■

7.9. Remark: (a) The same proof shows that 7.7 also holds with $X, \overline{X}, \overline{Y}$ replaced by countable families of relations of varying dimension.

(b) Natural generalizations of 7.7 hold with suitable hypothesis of κ-atomlessness, κ any uncountable cardinal, replacing $\aleph_1$-atomlessness. If we say that for $(\Omega, \mathcal{F}_n, P_n)_{n \in \mathbb{N}}$ a g.p.s., $(\mathcal{F}_n : n \in \mathbb{N})$ is κ-homogeneous if for each n $\mathcal{F}_n$ is κ-atomless and generated by κ sets over $\mathcal{F}_{<n}$, then one can easily prove a measure-algebra isomorphism theorem generalizing that of Maharam [12]; but, although such κ-homogeneous spaces do exist, this does not seem to lead simply to any classification theorem, as does Maharam's. ■

Now we can conveniently settle the question about when two representations may represent the same array.

7.10. Theorem: (a) Let $\{(U^t_\sigma : \phi \neq t \subseteq n)\}$ be an independent family of arrays. Then

$$\{f(U^t_\sigma : \phi \neq t \subseteq n)\} \sim \{g(U^t_\sigma : \phi \neq t \subseteq n)\}$$

iff there are measure preserving functions T

$$T^t : ([0,1]^2)^{2^t - \{\phi\}} \to [0,1] \qquad (2^t = \text{set of subsets of } t),\ \phi \neq t \subseteq n,$$

such that

(1) $\int \varphi(T^t(x^s, \bar{x}^s : \phi \neq s \subseteq t))dx^t \bar{x}^t$ is a.s. constant for every $\varphi \in C([0,1])$, $t \subseteq n$.

(2) $f(x^t : \phi \neq t \subseteq n) = g(T^t(x^s, \bar{x}^s : \phi \neq s \subseteq t) : t \subseteq n)$ a.s.

(b) If instead $\{(U^t_\sigma : t \subseteq n)\}$ is symmetric independent, then the same holds, with the extra condition that

(3) $T^{\pi(t)}(x^s, \bar{x}^s : s \subseteq \pi(t)) = T^t(x^{\pi(s)} \bar{x}^{\pi(s)} : s \subseteq t)$.

Proof: In both cases (<==) is straightforward: if $\{U^t_\sigma, \bar{U}^t_\sigma, \phi \neq t \subseteq n\}$ is independent or symmetric independent, then since each T^t is measure preserving,

$$V^t_\sigma = T(U^s_\sigma, \bar{U}^s_\sigma : s \subseteq t)$$

will be uniform, (1) quarantees that $\{V_\sigma^t\}_{\sigma \in \mathbb{N}^t}$, $t \subsetneq n$, are independent and each i.i.d., and in the DRE case condition (3) makes $V^t_{\sigma^\pi} = V_\sigma^{\pi(t)}$, $\pi \in n!$. Thus

$$\{f(U_\sigma^t : t \subseteq n)\} = \{g(V_\sigma^t : t \subseteq n)\} \sim \{g(U_\sigma^t : t \subseteq n)\} .$$

(==>) By Lemma 7.8 and Cor. 7.6 we can assume that we have independent (and in (b) symmetric) families of relations $\{U^t : t \subseteq n\}$, $\{V^t : t \subseteq n\}$ on a g.p.s. $(\Omega, \mathcal{F}_m, P_m)_{m \in \mathbb{N}}$ such that $\mathcal{F}_n$ is $\aleph_1$-atomless over $\mathcal{F}_{<n}$ and $f(U^t : t \subseteq n) = g(V^t : t \subseteq n)$. By Theorem 7.3(b), there are Borel-measureable functions T^t, $t \subseteq n$, and an independent (and in (b) symmetric) family $\{\overline{U}^t : t \subseteq n\}$ which is jointly independent with $\{U^t : t \subseteq n\}$, such that

$$T^t(U^s, \overline{U}^s : s \subseteq t) = V^t$$

(in (b) since

$$\begin{aligned} T^t(U^s, \overline{U}^s : s \subseteq t)^\pi &= T^t((U^s)^\pi, (\overline{U}^s)^\pi, \ s \subseteq t) \\ &= T^t(U^{\pi(s)}, \overline{U}^{\pi(s)} : s \subseteq t) \\ &= V^{\pi(t)}, \end{aligned}$$

we may assume $T^{\pi(t)}(x^s, \overline{x}^s : s \subseteq \pi(t)) = T^t(x^{\pi(s)}, \overline{x}^{\pi(s)} : s \subseteq t)$).

Now since $(U^s, \overline{U}^s : s \subseteq t)$ and V^t are both uniformly distributed on their respective ranges. T^t must be measure preserving; and since also

$$f(U^t : t \subseteq n) = g(T^t(U^s, \overline{U}^s : s \subseteq t) : t \subseteq n),$$

$$f(x^t : \phi \neq t \subseteq n) = g(T^t(x^s, \overline{x}^s : \phi \neq s \subseteq t) : \phi \neq t \subseteq n) \quad \text{a.s.}$$

with respect to Lebesgue measure on $([0,1]^2)^{2^n-1}$. And because V^t is independent of $\mathcal{F}_{<t}$, for every $\varphi \in C([0,1])$,

$$P(\varphi(V^t) \mid \mathcal{F}_{<t}) = \int \varphi(T(U^s, \bar{U}^s, x^t, \bar{x}^t : s \subseteq t) dx^t \bar{x}^t$$

is constant a. s., and this implies (1). ■

This result, too, generalizes straightforwardly to RCE and RE representations.

The criterion is simplest in the ED case when $n = 2$. Then, by 5.6(c), $\{f(U_i^1, U_j^2, U_{ij}^{12})\} \sim \{g(U_i^1 U_j^2, U_{ij}^{12})\}$ iff there are T_1, T_2 measure perserving such that

$$\int \varphi(f(x^1, x^2, x^{12}))dx^{12} = \int \varphi(g(T_1(x^1, \bar{x}^1), T_2(x^2, \bar{x}^2), x^{12}))dx^{12} \quad \text{a. s.}$$

for every bounded continuous φ.

References

1. Aldous, D. J., Representations for partially exchangeable arrays of random variables, to appear.
2. Chang, C. C. and H. J. Keisler, Model Theory, North Holland, 1974.
3. Eagleson, G. K. and N. C. Weber, Limit theorems for weakly exchangeable arrays, Proc. Cambridge Philos. Soc. 84(1979), 73-80.
4. Gaifman, H., Concerning measures in first order calculi, Israel J. Math. 2(1964), 1-18.
5. Hewitt, E. W. and L. J. Savage, Symmetric measures on cartesian products, Trans. Amer. Math. Soc. 80(1955), 470-501.
6. Hoover, D. N., Not Amer. Math. Soc. 25(1978), A-498 and A-600.
7. Hoover, D. N., A normal form theorem for the probability logic $L_{\omega_1 P}$, with applications, J. Symb. Logic, to appear.
8. Keisler, H. J., Hyperfinite model theory, in Gandy and Hyland eds., Logic Symposium '76, North Holland, 1977.
9. Krauss, P. H., Representations of symmetric probability models, J. Symb. Logic 34(1969), 183-193.
10. Loeb, P., Conversion from nonstandard to standard measure spaces and applications to probability, Trans. Amer. Math. Soc. 211(1975), 113-122.
11. Loeve, M., Probability Theory, Van Nostrand, 1960.
12. Maharam, D., On homogeneous measure algebras, Trans. Nat. Acad. Sci. 28(1942), 108-111.

13. Royden, H. L., Real Analysis, Macmillan, 1963.

14. Silverman, B. W., Limit theorems for dissociated random variables, Adv. Appl. Prob., 8(1976), 806-819.