FINITE AND INFINITE EXCHANGEABILITY

Takis Konstantopoulos
Uppsala University

5th Athens Probability Colloquium
Πανεπιστήμιο Αθηνών
20/5/2017

Joint work with Svante Janson and Linglong Yuan

PLAN

A bit of history and background

Infinite vs. finite extendibility

Examples

Urn measures

Homogeneous polynomials

Representation theorem

Extendibility

Background/History

Definition: A sequence $X = (X_1, X_2, \cdots)$ of random elements with values in some measurable space S is called exchangeable if its law is invariant under permutations of finitely many elements.

Concept first discussed by Jules Haag (ICM 1924)

The first rigorous representation theorem proved by de Finetti (1931).

Dynkin (1953) treated the case $S = \mathbb{R}$.

Hewitt and Savage (1955) generalized this to compact Hausdorff S equipped with the Baire σ-field.

Ryll-Nardzewski (1961) replaced exchangeability by the weaker notion of spreadability (any subsequence has the same law as the original sequence).

Dubins and Savage (1979) gave a counterexample to the representation theorem for a sufficiently weird space S.

The infinite exchangeability representation theorem

Theorem [de Finetti, Ryll-Nardzweski]. Suppose that S is a Borel space and $X = (X_1, X_2, \cdots)$ a spreadable sequence of random elements with values in S. Let $\mathscr{I}$ be the σ-field of invariant events. Consider the regular conditional probability

$$\eta(\omega, \cdot) := P(X_1 \in \cdot | \mathscr{I}),$$

as a random probability measure on S. Then

$$P(X \in \cdot | \mathscr{I}) = \eta^\infty,$$

where $\eta^\infty(\cdot, \omega)$ is the countably infinite product of $\eta(\cdot, \omega)$ by itself.

Corollary. In particular, for all measurable $A \subset S$,

$$P(X \in A) = E[\eta^\infty(A)] = \int_{\mathscr{P}(S)} \pi^\infty(A)\, \mu(d\pi)$$

where

$$\mathscr{P}(S) := \text{ set of probability measures on } S$$

(equipped with the standard σ-field induced by projections), and

$$\mu := \text{ law of } \eta, \quad \mu \in \mathscr{P}(\mathscr{P}(S)).$$

de Finetti follows from Birkhoff

The best proof is in Kallenberg's book on probabilistic symmetries. It goes like this.

By spreadability, for all $m \geq 1$,

$$(X_1, \ldots, X_{n-1}, X_n, X_{n+1}, \ldots) \stackrel{(d)}{=} (X_1, \ldots, X_{n-1}, X_{n+m}, X_{n+m+1}, \ldots)$$

and so for $f_1, \ldots, f_n$ bounded measurable on S and g bounded and shift-invariant on S^∞,

$$E[f_1(X_1) \cdots f_n(X_n) g(X)] = E[f_1(X_1) \cdots f_{n-1}(X_{n-1})\, R_{m,n}\, g(X)]$$

where $R_{m,n} = m^{-1} \sum_{j=1}^m f_n(X_{n+j}) \to \eta[f_n]$, as $m \to \infty$, a.s., by the ergodic theorem. Therefore,

$$\begin{aligned} E[f_1(X_1) \cdots f_n(X_n) g(X)] &= E[f_1(X_1) \cdots f_{n-1}(X_{n-1})\, \eta[f_n]\, g(X)] \\ \cdots &= E[\eta[f_1] \cdots \eta[f_{n-1}]\, \eta[f_n]\, g(X)] \end{aligned}$$

and this obviously implies that

$$P((X_1, \ldots, X_n) \in \cdot | \mathscr{I}) = \eta^n,$$

from which the result is immediate.

N.B.1. The assumption that S is Borel is only needed to ensure that the regular conditional distribution η exists.

N.B.2. The probability measure μ in the representation theorem is unique.

N.B.3. Obviously, the representation theorem also holds for arbitrary Cartesian products S^T rather than $S^{\mathbb{N}}$ of a Borel space S.

My favorite example

Let $B = (B(t), t \in \mathbb{R})$ be standard Brownian motion with 2-sided parameter. Let $B_1, B_2, \ldots$ be i.i.d. copies of B. Then

$$W_n := B_n \circ \cdots \circ B_1 \to W_\infty,$$

in the sense of convergence of finite-dimensional distributions. Furthermore,

$$(W_\infty(t), \quad t \neq 0)$$

is exchangeable and, therefore, by de Finetti's theorem, a mixture of i.i.d. random variables. The mixing measure μ can be found as follows. Let

$$\eta_n := \int_0^1 \mathbf{1}\{W_n(t) \in \cdot\}\, dt$$

be the occupation measure of W_n on the unit interval. Then η_n converges, in distribution, as a random element of $\mathscr{P}(\mathbb{R})$ to a random probability measure η_∞. Then μ is the law of η_∞.

C.F. Curien and K. (2012). Iterating BMs ad libitum. JOTP 27, 433-448.

Finite exchangeability

$(X_1, \dots, X_n)$ is (finitely or n-)exchangeable if its law is invariant under all $n!$ permutations.

Finite exchangeability is often more natural than infinite exchangeability.

Examples:

1) In Statistics, one may deal with unordered data but the size is always finite. Therefore, the assumption that the data comes from an exchangeable infinite sequence may be impractical or wrong.

2) Draw n balls at random from an urn containing $N \geq n$ balls, some of which are colored red, some blue, etc.

3) The n-coalescent.

4) The Curie-Weiss Ising model in n dimensions.

5) A random vector in $\mathbb{R}^n$ with density being a symmetric function.

Natural questions

1. Does a de Finetti-type representation result hold?

2. Given an n-exchangeable sequence $(X_1, \ldots, X_n)$ and $N > n$ is there an N-exchangeable sequence $(Y_1, \ldots, Y_N)$ such that $(X_1, \ldots, X_n)$ has the same law as $(Y_1, \ldots, Y_n)$?

Answers are no and no, in general.

Example. An urn contains one red and one blue ball. Pick them at random.

The probability measure $P = \frac{1}{2}\delta_{rb} + \frac{1}{2}\delta_{br}$ on $\{r, b\}^2$ cannot be written as a mixture of independent measures.

Moreover, there is no exchangeable probability measure Q on $\{r, b\}^3$ such that its projection of $\{r, b\}^2$ be equal to P.

Finite exchangeability representation result

Theorem. Let $(X_1, \ldots, X_n)$ be an exchangeable sequence of length n of random elements of an *arbitrary* measurable space S. Then there is a finite signed measure ξ on $\mathscr{P}(S)$ such that

$$P((X_1, \ldots, X_n) \in A) = \int_{\mathscr{P}(S)} \pi^n(A)\xi(d\pi),$$

for measurable $A \subset S^n$.

References

– Jaynes (1986).

– Diaconis (1977).

– Kerns and Szekely (2006).

– Janson, K. and Yuan (2016).

An algebraic result

Let n and d be positive integers. A composition of length n of d is a sequence $\lambda = (\lambda_1, \ldots, \lambda_n)$ of n nonnegative integers such that $\lambda_1 + \cdots + \lambda_d = n$. The number of such compositions is the number of placements of n unlabelled balls in d labelled boxes, that is, $\binom{n+d-1}{d-1}$.

Denote by $\mathcal{N}_n(d)$ the set of the n-compositions of d.

Theorem. The polynomials

$$p_\lambda(x_1, \ldots, x_d) := (\lambda_1 x_1 + \cdots + \lambda_d x_d)^n, \quad \lambda \in \mathcal{N}_n(d),$$

form a basis of the space of all homogeneous polynomials of degree n in d variables $x_1, \ldots, x_d$.

N.B.1. It is obvious that $x_1^{\lambda_1} \cdots x_d^{\lambda_d}$, $\lambda \in \mathcal{N}_n(d)$, form a basis for the space of degree-n homogeneous polynomials in d variables but this is not of immediate help.

N.B.2. The theorem above is equivalent to the statement that the multinomial Dyson matrix (the transition probability matrix of the Wright-Fisher Markov chain with d types) is invertible.

Urn measures-idea of proof

Let σ be a permutation of $\{1, \dots, n\}$. With $\sigma X = (X_{\sigma(1)}, \dots, X_{\sigma(n)})$, we have

$$P(X \in A) = \frac{1}{n!} \sum_{\sigma} P(\sigma X \in A) = EU_X(A),$$

where

$$U_x = \frac{1}{n!} \sum_{\sigma} \delta_{\sigma x}$$

is an urn measure: an urn contains items labelled $x_1, \dots, x_n$; select all of them, without replacement, at random. The reason for the representation is that U_X itself can be written as an integral with respect to a random signed measure on the space of probability measures of S.

We explain this for the case $|S| = d < \infty$. The general case is a bit more involved.

A point measure ν on S is a measure with nonnegative integer values. Let $\mathcal{N}_n(S)$ be the set of point measures of total mass n. For $x \in S^n$ let $\varepsilon_x = \sum_{i=1}^n \delta_{x_i} \in \mathcal{N}_n(S)$. Let $S^n(\nu) := \{x \in S^n : \varepsilon_x = \nu\}$. Then

$$S^n = \bigcup_{\nu \in \mathcal{N}_n(S)} S^n(\nu),$$

and the union is disjoint. Note $|S^n(\nu)| = \binom{n}{\nu} = n!/\prod_a \nu\{a\}!$.

...idea of proof

$S^n(\nu)$ is an "urn". There is only one exchangeable probability measure supported on $S^n(\nu)$, the uniform measure:

$$u_\nu = \binom{n}{\nu}^{-1} \sum_{z \in S^n(\nu)} \delta_z.$$

Hence, if Q is exchangeable probability measure on S^n we have

$$Q = \sum_{\nu \in \mathcal{N}_n(S)} Q(S^n(\nu))\, u_\nu.$$

In particular, with

$$Q = \pi^\nu, \quad \pi \in \mathscr{P}(S),$$

we have $Q(T^n(\nu)) = \binom{n}{\nu}\pi^\nu = \binom{n}{\nu}\prod_a \pi\{a\}^{\nu\{a\}}$. Specializing further, take

$$\pi = \frac{1}{n}\lambda, \quad \lambda \in \mathcal{N}_n(S).$$

Then

$$\lambda^n = \sum_{\nu \in \mathcal{N}_n(S)} \binom{n}{\nu} \lambda^\nu\, u_\nu.$$

The earlier algebraic result says that

$$u_\nu = \sum_{\lambda \in \mathcal{N}_n(S)} M(\nu, \lambda)\, \lambda^n.$$

On noticing that $U_x = u_{\varepsilon_x}$ we complete the proof for the finite S case.

Extendibility

Theorem. Let S be a lcH (locally compact Hausdorff space) and $(X_1, \ldots, X_n)$ a random element of S^n with exchangeable law such that the law of X_1 is inner and outer regular. Let $N > n$. Then $(X_1, \ldots, X_n)$ is N-extendible if and only if, for all $\varepsilon > 0$ and all bounded measurable $f : S^n \to \mathbb{R}$, there is $(a_1, \ldots, a_N) \in S^N$ such that

$$|Ef(X_1, \ldots, X_n)| \leq \frac{1+\varepsilon}{N(N-1)\cdots(N-n+1)} \left| \sum_\sigma f(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) \right|$$

where the sum ranges over all one-to-one functions $\sigma : \{1, \ldots, n\} \to \{1, \ldots, N\}$.

C.F. K and Yuan (2016).

Symmetrizing operators

The idea of proof is based on the following.

Let $f(x_1, \dots, x_n)$ be a real-valued function of n variables. We can create a *symmetric* function of N variables by

$$U_n^N f(x_1 \dots, x_N) = \frac{1}{(N)_n} \sum f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

where $(N)_n = N(N-1) \cdots (N-n+1)$.

Probabilistically, we select at random, without replacement, n items from an urn containing the items $x_1, \dots, x_N$ and evaluate f at the selected items.

Let $b(S^n)$ be the space of bounded measurable real-valued functions on S^n, equipped with the sup norm and let $U_n^N b(S_n)$ be its image under U_n^N. We next define the linear functional

$$\mathcal{E} : U_n^N b(S^n) \to \mathbb{R}$$

by the recipe

$$\mathcal{E}(U_n^N f) = Ef(X_1, \dots, X_n).$$

It is not clear that $\mathcal{E}$ is a function. But it is, due to algebraic properties of urn measures.

If $(X_1, \dots, X_n)$ is exchangeable then $\|\mathcal{E}\| = 1$ and this translates to the condition of the theorem. The converse needs work.

The converse

Suppose $\|\mathcal{E}\| = 1$. The idea for proving the bulk of the theorem is based on

Extend $\mathcal{E}$ to $b_{sym}(S^N)$ via the Hahn-Banach theorem and let $\mathcal{E}'$ be the extension. Let $\mathcal{L} : b(S^N) \to \mathbb{R}$ be obtained by symmetrization:

$$\mathcal{L} = \mathcal{E}' \circ U_N^N.$$

We have

$$\|\mathcal{L}\| = \|\mathcal{E}'\| = \|\mathcal{E}\|.$$

This implies that $F(A) := \mathcal{L}(\mathbf{1}_A)$ is a finitely additive nonnegative set function. But it is not a probability measure even in "good" cases. The point is to extract a probability measure.

Next, restrict $\mathcal{L}$ onto the space $C_c(S^N)$ of continuous compactly supported real-valued functions on S^N and use Urysohn's lemma and inner regularity of the law of $(X_1, \ldots, X_n)$ to deduce that $\|\mathcal{L}|C_c(S^N)\| = 1$.

Then use the Riesz representation theorem and outer regularity in order to represent

$$\mathcal{L}f = \int_{S^N} f d\lambda$$

for some symmetric regular probability measure λ on S^N.

It turns out that a random element $(Y_1, \ldots, Y_N)$ of S^N with law λ is an N-extension of $(X_1, \ldots, X_n)$.

Recent references

Svante Janson, T.K. and Linglong Yuan,
A direct proof of the finite exchangeability representation theorem.
J. Math. Anal. Appl. 442, no. 2, 703-714, 2016.

T.K. and Linglong Yuan,
On the extendibility of finitely exchangeable probability measures.
arXiv:1501.06188 [math.PR]