EXCHANGEABLE SEQUENCES AND DE FINETTI'S THEOREM

In these notes we summarize the lecture about de Finetti's Theorem, concerning *exchangeable sequences*. Here $\mathbf{X} = (X_1, X_2, \dots)$ will denote a sequence of random variables X_i *with values in* $\{0, 1\}$. (The things discussed here extend also to X_i with values in $\mathbb{R}$, or even more general spaces).

DEFINITION. *The sequence* $\mathbf{X}$ *is called* exchangeable *if for all* n *and all permutations* π *of* $1, 2, \dots, n$ *we have that*

$$(X_1, X_2, X_3, \dots) \stackrel{\mathrm{d}}{=} (X_{\pi(1)}, X_{\pi(2)}, X_{\pi(3)}, \dots).$$

Note that we only consider permutations π which are 'finite' in the sense that they permute finitely many (n) numbers.

EXERCISE 1. *Prove that* $\mathbf{X}$ *is exchangeable if and only if for each* $n > 1$,

$$\mathbf{X} \stackrel{\mathrm{d}}{=} (X_n, X_2, X_3, \dots, X_{n-1}, X_1, X_{n+1}, X_{n+2}, \dots)$$

Prove that $\mathbf{X}$ *is exchangeable if and only if for each bijection* φ *from* $\{1, 2, 3, \dots\}$ *to itself,*

$$\mathbf{X} \stackrel{\mathrm{d}}{=} (X_{\varphi(1)}, X_{\varphi(2)}, X_{\varphi(3)}, \dots).$$

One example of an exchangeable sequence is an i.i.d. sequence. Another example is a 'constant' sequence, $X_i = X$ for all i and some random variable X. An exchangeable sequence is always strongly stationary.

EXERCISE 2. *Prove that the set of exchangeable sequences is* convex *in the following sense: if* $\mathbf{X}$ *and* $\mathbf{X}'$ *are exchangeable, then the sequence given by*

$$\begin{cases} \mathbf{X}, & \text{with probability } p, \\ \mathbf{X}', & \text{with probability } 1-p, \end{cases}$$

is also exchangeable, for any $p \in [0, 1]$.

Recall that the law of any sequence $\mathbf{X}$ of random variables in $\{0, 1\}$ (exchangeable or not) can be described in terms of a probability measure ν on the sequences $\{0, 1\}^{\mathbb{N}}$. It is a fact of measure theory, which we don't prove, that such a probability measure is determined by its values on *cylinder sets*, by which we mean the following. Given S_0 and S_1, disjoint subsets of $\{1, 2, 3, \dots\}$, let

$$C(S_0, S_1) = \big\{\mathbf{x} \in \{0, 1\}^{\mathbb{N}} : x_i = 0 \text{ for all } i \in S_0, x_i = 1 \text{ for all } i \in S_1\big\},$$

and let $C_{k,\ell} = C(\{1, \dots, k\}, \{k + 1, \dots, k + \ell\})$.

THEOREM (de Finetti's theorem). *If* $\mathbf{X}$ *is exchangeable, and* ν *is the corresponding probability measure on sequences, then there is some random variable* $\xi \in [0, 1]$ *such that*

$$\nu(C_{k,\ell}) = \mathbb{E}\big[(1 - \xi)^k \xi^\ell\big].$$

The interpretation is as follows. Inside the $\mathbb{E}[\cdots]$, the expression $(1 - \xi)^k \xi^\ell$ is what we would get from an i.i.d. Ber(ξ) sequence. So what we have is that $\mathbf{X}$ is like an i.i.d. Bernoulli sequence with a *random* probability (ξ) for 1. This means that an echangeable sequence is a 'generalized convex combination' of i.i.d. sequences.

A key part of the proof is the following consequence of the Stone–Weierstrass theorem:

LEMMA. *If* ξ_n *are random variables on* $[0, 1]$ *then there is a subsequence which converges in distribution to a limit* ξ.

Proof of de Finetti's Theorem. Writing $P(k,\ell) = \nu(C_{k,\ell})$ we have that $P(k,\ell) = P(k,\ell+1) + P(k+1,\ell)$. Iterating this, for any fixed k_0, ℓ_0 and any $n \geq n_0 = k_0 + \ell_0$

$$P(k,\ell) = \sum_{k+\ell=n} \frac{\binom{n-n_0}{k-k_0}}{\binom{n}{k}} M^{(n)}(k,\ell), \quad \text{where } M^{(n)}(k,\ell) = \binom{n}{k} P(k,\ell).$$

Now the $M^{(n)}(\ell, n-\ell)$ sum to 1 as ℓ ranges from 0 to n, so we can define random variables ξ_n with values between 0 and 1, by

$$\mathbb{P}(\xi_n = \ell/n) = M^{(n)}(n-\ell,\ell).$$

Note that

$$\mathbb{E}\big[(1-\xi_n)^{k_0}\xi_n^{\ell_0}\big] = \sum_{k+\ell=n} \big(\tfrac{k}{n}\big)^{k_0}\big(\tfrac{\ell}{n}\big)^{\ell_0} M^{(n)}(k,\ell).$$

By extracting a subsequence if necessary (using the Lemma) we can assume that ξ_n converge in distribution to some ξ. Then, since the function $f(x) = (1-x)^{k_0}x^{\ell_0}$ is continuous, the result follows from the estimate given in the exercise below. ☐

Exercise 3. *For $k_0, \ell_0, n_0, k, \ell, n$ as above, there is a constant $C(k_0,\ell_0)$ such that*

$$\left| \frac{\binom{n-n_0}{k-k_0}}{\binom{n}{k}} - \big(\tfrac{k}{n}\big)^{k_0}\big(\tfrac{\ell}{n}\big)^{\ell_0} \right| \leq \tfrac{C(k_0,\ell_0)}{n}.$$

Exercise 4. *Let $\mathbf{X}$ be an exchangeable sequence. Prove that* $\operatorname{Cov}(X_i, X_j) \geq 0$ *for all i and j.*

Exercise 5. *What is the appropriate generalization of de Finetti's Theorem if the X_i take values in the set $\{1, 2, \dots, N\}$ for some $N \geq 2$? Can you adapt the proof above?*

Exercise 6. *Recall that an exchangeable sequence is strongly stationary. In this exercise we apply the ergodic theorem to identify the limit of $\frac{1}{n}\sum_{i=1}^n X_i$ for an exchangeable sequence $\mathbf{X} \in \{0,1\}^{\mathbb{N}}$, using the result of Exercise 9.7.13 in Grimmett–Stirzaker.*

The set-up is as follows. Let $F(t)$ be the cumulative distribution function for the ξ in de Finetti's theorem, and

- $\Omega = [0,1] \times [0,1]^{\mathbb{N}}$ *where $\omega \in \Omega$ is written as $\omega = (\omega_0; \omega_1, \omega_2, \dots)$,*
- $\tau : \Omega \to \Omega$ *is given by $\tau(\omega_0; \omega_1, \omega_2, \dots) = (\omega_0; \omega_2, \omega_3, \dots)$,*
- $\mathbb{P}$ *is given by*

$$\mathbb{P}\big((a_0,b_0] \times (a_1,b_1] \times \cdots \times (a_n,b_n] \times [0,1]^{\mathbb{N}}\big) = \big(F(b_0) - F(a_0)\big)\big(b_1 - a_1\big)\big(b_2 - a_2\big)\cdots\big(b_n - a_n\big),$$

- $\xi(\omega) = \omega_0$, $U(\omega) = \omega_1$, $X(\omega) = \mathbb{1}\{\omega_1 \leq \omega_0\}$, $U_i(\omega) = U(\tau^{i-1}(\omega))$, $X_i(\omega) = X(\tau^{i-1}(\omega))$

(a) *Show that, with this definition, we indeed have $\mathbb{P}(\xi \leq t) = F(t)$, and also that the U_i are independent U[0,1] random variables.*

(b) *Show that $\mathbf{X} = (X_1, X_2, \dots)$ is an exchangeable sequence satisfying*

$$\mathbb{P}(\mathbf{X} \in C_{k,\ell}) = \mathbb{E}[(1-\xi)^k \xi^\ell].$$

(c) *Show (using the definition of conditional expectation) that $\mathbb{E}[X \mid \mathcal{I}] = \xi$, and thus*

$$\frac{1}{n}\sum_{i=1}^n X_i \to \xi \qquad \text{a.s. and in } L^1.$$