Math 280C, Spring 2005

Exchangeability

Let $X = (X_1, X_2, \ldots)$ be an exchangeable sequence of random variables. As discussed in class, there is no loss of generality (and some gain of convenience) in assuming that the sample space Ω is the sequence space $\mathbf{R}^{\mathbf{N}}$ (with generic element $\omega = (\omega_1, \omega_2, \ldots)$) endowed with the product σ-field $\mathcal{F} := \mathcal{B}^{\mathbf{N}}$ ($\mathcal{B}$ the Borel σ-field on $\mathbf{R}$), and that X is the sequence of coordinate random variables:

$$X_n(\omega) := \omega_n, \qquad n = 1, 2, \ldots.$$

Notice that $\mathcal{F} = \sigma\{X_1, X_2, \ldots\}$.

The tail σ-field $\mathcal{T}$ is defined as

$$\mathcal{T} := \cap_n \mathcal{T}_n,$$

where

$$\mathcal{T}_n := \sigma\{X_{n+1}, X_{n+2}, \ldots\}, \qquad n = 1, 2, \ldots.$$

Let Σ_n denote the set of all permutations σ of $\mathbf{N}$ with the property that $\sigma(k) = k$ for all $k > n$, and put $\Sigma = \cup_n \Sigma_n$, the set of all finite permutations of $\mathbf{N}$. Recall that the exchangeability of X means that

$$\sigma X := (X_{\sigma(1)}, X_{\sigma(2)}, \ldots)$$

has the same distribution as X for each $\sigma \in \Sigma$. Put another way, we can view $\sigma \in \Sigma$ as inducing a mapping (also called σ) of Ω onto itself:

$$\sigma(\omega) := (\omega_{\sigma(1)}, \omega_{\sigma(2)}, \ldots), \qquad \omega \in \Omega,$$

and then exchangeability means that $\mathbf{P}[\sigma^{-1}(A)] = \mathbf{P}[A]$ for each $\sigma \in \Sigma$ and each $A \in \mathcal{F}$. (Here $\mathbf{P}$ is the probability measure on $(\Omega, \mathcal{F})$ governing X.) If $A \in \mathcal{F}$, then we say that A is *n-symmetric* provided $\sigma^{-1}A = A$ for every $\sigma \in \Sigma_n$. It is easy to check that the collection of such events forms a σ-field, which we denote $\mathcal{E}_n$. Evidently $\mathcal{E}_n \supset \mathcal{E}_{n+1}$; with this in mind we define the *exchangeable* σ-field $\mathcal{E}$ as

$$\mathcal{E} := \cap_n \mathcal{E}_n.$$

Of course, $\mathcal{E}$ is just the class of all events $A \in \mathcal{F}$ such that $\sigma^{-1}(A) = A$ for all $\sigma \in \Sigma$. Notice that

$$\mathcal{T}_n \subset \mathcal{E}_n, \qquad \forall n \in \mathbf{N},$$

whence

$$\mathcal{T} \subset \mathcal{E}.$$

This inclusion is strict; for example, the event

$$A := \left\{ \omega : \sum_{k=1}^{n} X_k(\omega) \in [0,1] \text{ for infinitely many } n \in \mathbf{N} \right\}.$$

lies in $\mathcal{E}$ but not in $\mathcal{T}$. (Explain!) Nonetheless, for exchangeable sequences there is not much of a difference between $\mathcal{T}$ and $\mathcal{E}$, as the following result of P.-A. Meyer shows. Its corollary, the Hewitt-Savage 0-1 law, originates in [**1**].

Theorem 1. *If $A \in \mathcal{E}$, then there exists $B \in \mathcal{T}$ such that $\mathbf{P}[A \triangle B] = 0$.*

Proof. Define $\mathcal{F}_n := \sigma\{X_1, X_2, \ldots, X_n\}$, and $Z := 1_A$. By the martingale convergence theorems (forward and reverse):

$$\lim_n \mathbf{E}[Z|\mathcal{F}_n] = Z,$$

and

$$\lim_n \mathbf{E}[Z|\mathcal{T}_n] = \mathbf{E}[Z|\mathcal{T}].$$

in L^1 (and a.s.). Thus, given $\epsilon > 0$ there exists n_0 so large that

$$\|Z - \mathbf{E}[Z|\mathcal{F}_n]\|_1 < \epsilon \tag{1}$$

and

$$\|\mathbf{E}[Z|\mathcal{T}] - \mathbf{E}[Z|\mathcal{T}_n]\|_1 < \epsilon \tag{2}$$

provided $n \geq n_0$. For such an n let σ be the permutation that exchanges 1 and $n+1$, 2 and $n+2$, ..., n and $2n$, and leaves the other integers alone. Because $A \in \mathcal{E}$, we have $Z \circ \sigma = Z$; because of this and the exchangeability of X, (1) implies

$$\|Z - \mathbf{E}[Z|\sigma\{X_{n+1}, \ldots, X_{2n}\}]\|_1 < \epsilon. \tag{3}$$

But then, by the tower property and the L^1-contraction property of conditional expectation,

$$\begin{aligned} \|\mathbf{E}[Z|\mathcal{T}_n] - \mathbf{E}[Z|\sigma\{X_{n+1}, \ldots, X_{2n}\}]\|_1 &= \left\| \mathbf{E}\left[Z - \mathbf{E}[Z|\sigma\{X_{n+1}, \ldots, X_{2n}\}] \middle| \mathcal{T}_n \right] \right\|_1 \\ &\leq \|Z - \mathbf{E}[Z|\sigma\{X_{n+1}, \ldots, X_{2n}\}]\|_1 < \epsilon \end{aligned} \tag{4}$$

Combining (2), (3), and (4), we obtain

$$\|Z - \mathbf{E}[Z|\mathcal{T}]\|_1 < 3\epsilon.$$

Since $\epsilon > 0$ was arbitrary, it follows that

(5) $$Z = \mathbf{E}[Z|\mathcal{T}] \text{ almost surely.}$$

The event $B := \{\omega : \mathbf{E}[Z|\mathcal{T}](\omega) = 1\}$ is $\mathcal{T}$-measurable, and (5) means that $\mathbf{P}[A \triangle B] = 0$. ∎

Corollary. [Hewitt-Savage 0-1 law] *If the exchangeable sequence X is in fact iid, then $\mathcal{E}$ is trivial: $\mathbf{P}[A] = 0$ or 1 for every $A \in \mathcal{E}$.*

Proof. By Kolmogorov's 0-1 law, the tail σ-field $\mathcal{T}$ is trivial if the X_ks are independent. ∎

The following result was mentioned in class. The general case, in which the random variables $X_1, X_2, \ldots$ are permitted to take values in a space more general than the real line, is due to E. Hewitt and L.J. Savage; see [**1**]. A splendid introduction to this result, and its applications, can be found in [**2**].

Theorem 2. [De Finetti's Theorem] *The sequence $X = (X_1, X_2, \ldots)$ is conditionally iid, given $\mathcal{E}$. More precisely, if $f_1, f_2, \ldots, f_k$ are bounded and measurable, then*

$$\mathbf{E}[f_1(X_1)f_2(X_2)\cdots f_k(X_k)|\mathcal{E}] = \prod_{j=1}^{k} \mathbf{E}[f_j(X_1)|\mathcal{E}].$$

Proof. I shall write out the proof only for $k = 2$. The general case is not essentially different. By a symmetry argument used in class, we have (for $n \geq 2$)

(6) $$\mathbf{E}[f_1(X_1)f_2(X_2)|\mathcal{E}_n] = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} f_1(X_i)f_2(X_j).$$

Now the right side of (6) differs from

$$\left\{ \frac{1}{n} \sum_{i=1}^{n} f_1(X_i) \right\} \left\{ \frac{1}{n} \sum_{j=1}^{n} f_2(X_j) \right\}$$

only in that the latter contains "diagonal" terms and the former does not. In fact, these two differ by no more that $2\|f\|_\infty^2/n$. It follows that, almost surely,

$$\begin{aligned}
\mathbf{E}[f_1(X_1)f_2(X_2)|\mathcal{E}] &= \lim_n \mathbf{E}[f_1(X_1)f_2(X_2)|\mathcal{E}_n] = \lim_n \left\{ \frac{1}{n} \sum_{i=1}^{n} f_1(X_i) \right\} \left\{ \frac{1}{n} \sum_{j=1}^{n} f_2(X_j) \right\} \\
&= \lim_n \mathbf{E}[f_1(X_1)|\mathcal{E}_n] \cdot \mathbf{E}[f_2(X_1)|\mathcal{E}_n] \\
&= \mathbf{E}[f(X_1)|\mathcal{E}] \cdot \mathbf{E}[f_2(X_1)|\mathcal{E}],
\end{aligned}$$

as claimed. ∎

References

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

[2] J.F.C. Kingman: Uses of exchangeability. *Ann. Probability* **6** (1978) 183–197.