Lecture-28: Exchangeability

1 Exchangeability

Definition 1.1. A *finite permutation* of $\mathbb{N}$ is a bijective map $\alpha : \mathbb{N} \to \mathbb{N}$ such that $\alpha(i) \neq i$ for only finitely many i. That is, for a finite subset $F \subset \mathbb{N}$, we have $\alpha(F) = \{\alpha(i) : i \in F\} = F$ and $\alpha(i) = i$ for $i \notin F$.

Remark 1. It is clear that a finite permutation α can always be defined on an interval of form $[n]$, where $n = \max\{i \in \mathbb{N} : \alpha(i) \neq i\}$.

Definition 1.2. We define a projection operator $\pi_i : \prod_{n\in\mathbb{N}} \Omega_n \to \Omega_i$ such that $\pi_i(\omega) \triangleq \omega_i$ for any sequence $\omega \in \prod_{n\in\mathbb{N}} \Omega_n$.

Definition 1.3. Let $X_i : \Omega_i \to \mathcal{X}$ be a random variable on the probability space $(\Omega_i, \mathcal{S}_i, \mu_i)$. Consider the probability space $(\Omega, \mathcal{F}, P)$ for the process $X : \Omega \to \mathcal{X}^{\mathbb{N}}$, where

$$\Omega = \Omega_1 \times \Omega_2 \times \dots, \qquad \mathcal{F} = \mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \dots, \qquad P = \mu_1 \otimes \mu_2 \otimes \dots.$$

Remark 2. For a projection operation $\pi_i : \prod_{n\in\mathbb{N}} \Omega_n \to \Omega_i$ and any event $A_i \in \mathcal{S}_i$, we have

$$\pi_i^{-1}(A_i) = \Omega_1 \times \dots \times A_i \times \dots \in \mathcal{F}.$$

This also implies that $(P \circ \pi_i^{-1})(A_i) = \mu_i(A_i)$ for any $A_i \in \mathcal{S}_i$ and hence $\mu_i = P \circ \pi_i^{-1}$.

Definition 1.4. Consider an outcome $\omega \in \Omega \triangleq \prod_{n\in\mathbb{N}} \Omega_n$, the projection operator $\pi_i : \Omega \to \Omega_i$ for some $i \in \mathbb{N}$, and a finite permutation $\alpha : \mathbb{N} \to \mathbb{N}$, then we can define a finitely permuted outcome $\alpha(\omega) \triangleq (\omega_{\alpha(i)} : i \in \mathbb{N})$ in terms of its projections for each outcome $\omega \in \Omega$, as

$$(\pi_i \circ \alpha)(\omega) \triangleq \pi_{\alpha(i)}(\omega).$$

Definition 1.5. An event $A \in \mathcal{F}$ is *n-permutable* if for all n-permutations $\alpha : [n] \to [n]$, we have

$$A = \alpha^{-1}(A) = \{\omega \in \Omega : \alpha(\omega) \in A\}.$$

An event $A \in \mathcal{F}$ is *permutable* if it is n-permutable for all $n \in \mathbb{N}$.

Example 1.6 (Permutable event). Consider a random sequence $\omega \in \Omega = \{H, T\}^{\mathbb{N}}$, then the event $A \triangleq \{k \text{ heads in first } n \text{ tosses}\}$ is n-permutable.

Example 1.7 (Non-permutable event). Consider a random sequence $\omega \in \Omega = \{H, T\}^{\mathbb{N}}$, then the event $A \triangleq \{\omega_1 = H, \omega_2 = T\}$ is not permutable.

Example 1.8 (One-dimensional random walk). Consider a one-dimensional random walk $S_n \triangleq \sum_{i=1}^n X_i, \quad n \in \mathbb{N}$ defined for *i.i.d.* step-size sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$. Let $c \in (\mathbb{R} \setminus \{0\})^{\mathbb{N}}$ be a non-zero constant sequence. We define the following events

$$A \triangleq \cap_{n\in\mathbb{N}} \cup_{k\geqslant n} \{S_k \leqslant x\} = \{S_n \leqslant x \text{ infinitely often }\}, \qquad B \triangleq \left\{ \limsup_{n\in\mathbb{N}} \frac{S_n}{c_n} \geqslant 1 \right\}.$$

The events A, B are permutable. This is due to the fact that $S_n(X) = S_n(\alpha(X))$ for any n-permutation α, and hence A and B are n-permutable.

Example 1.9 (Tail σ-algebra). For any random sequence $X : \Omega \to \mathcal{X}^{\mathbb{N}}$, we define $\mathcal{G}_n \triangleq \alpha(X_k, k \geqslant n)$ for each $n \in \mathbb{N}$ and tail σ-algebra as $\mathcal{T} \triangleq \cap_{n\in\mathbb{N}} \mathcal{G}_n$. Then, all events in $\mathcal{T}$ are permutable. This follows from the fact that any event $A \in \mathcal{T}$ belongs to $\mathcal{G}_{n+1}$ for all $n \in \mathbb{N}$. Further, any $A \in \mathcal{G}_n$ remains unaffected by the permutation of $(X_1, \dots, X_n)$ and hence A is n-permutable.

Definition 1.10. The collection of all n-permutable events is a σ-algebra called *n-exchangeable* and is denoted by $\mathcal{E}_n$. The collection of permutable events is a σ-algebra called *exchangeable* and is denoted by $\mathcal{E}$.

Definition 1.11. A random sequence $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ is called *exchangeable* if for each n-permutation $\alpha : [n] \to [n]$, the joint distribution of $(X_1, X_2, \dots, X_n)$ and $(X_{\alpha(1)}, X_{\alpha(2)}, \dots, X_{\alpha(n)})$ are identical.

Remark 3. Observe that permutable is measure-independent, while exchangability is measure-dependent.

Remark 4. A random process $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ is exchangeable if all the events in its event space are permutable. That is, $\sigma(X) \subseteq \mathcal{E}$.

Example 1.12 (Draw without replacement). Suppose balls are selected uniformly at random, without replacement, from an urn consisting of n balls of which k are white. For draw $i \in [n]$, let ξ_i be the indicator of the event that the ith selection is white. Then the finite collection $(\xi_1, \dots, \xi_n)$ is exchangeable but not independent. In particular, we consider the random index set $W \triangleq \{i \in [n] : \xi_i = 1\}$. It follows that $|W| = k$, and we can write the probability of the event $\{W = A\} \in \mathcal{F}$ for some index set $A \subseteq [n]$ such that $|A| = k$ as

$$P\{W = A\} = P\{\xi_i = 1, i \in A, \xi_j = 0, j \notin A\} = \frac{k \dots 1 \times (n-k) \dots 1}{n \dots 1} = \frac{(n-k)!k!}{n!} = \frac{1}{\binom{n}{k}}.$$

This joint distribution is independent of set of exact locations A, and hence exchangeable. In addition, one can see the dependence from

$$P(\{\xi_2 = 1\} \mid \{\xi_1 = 1\}) = \frac{k-1}{n-1} \neq \frac{k}{n-1} = P(\{\xi_2 = 1\} \mid \{\xi_1 = 0\}).$$

Example 1.13 (Conditionally independent sequence). Consider a finite set $\mathcal{Y}$ and a random variable $Y : \Omega \to \mathcal{Y}$ with probability mass function $p \in \mathcal{M}(\mathcal{Y})$. Let $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ be a conditionally *i.i.d.* random sequence given random variable Y, with conditional distribution $F_{X|y} \triangleq F_{X|\{Y=y\}}$. We can write the joint finite dimensional distribution of the sequence X,

$$P\Big(\cap_{i=1}^n \{X_i \leqslant x_i\}\Big) = \sum_{y \in \mathcal{Y}} P\Big(\cap_{i=1}^n \{X_i \leqslant x_i\} \mid \{Y = y\}\Big) P\{Y = y\} = \sum_{y \in \mathcal{Y}} p_y \prod_{i=1}^n F_{X|y}(x_i)$$

Since any finite dimensional distribution of the sequence X is symmetric in $(x_1, \dots x_n)$, it follows that X is exchangeable.

Definition 1.14. Let $I_{n,k} \triangleq \Big\{i \in [n]^k : i_j \text{ distinct}\Big\}$. Then, the cardinality of this set is denoted by

$$(n)_k \triangleq |I_{n,k}| = n(n-1)\dots(n-k+1) = \binom{n}{k} k!.$$

Theorem 1.15 (De Finetti's Theorem). *If random sequence $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ is exchangeable, then the sequence X is* i.i.d. *conditioned on exchangeable σ-algebra $\mathcal{E} \triangleq \sigma(X)$.*

Proof. To show the independence of exchangeable random sequence X, conditioned on exchangeable σ-algebra $\mathcal{E}$, it suffices to show that for bounded functions $f_i : \mathbb{R} \to \mathbb{R}$ for each i, we have

$$\mathbb{E}[\prod_{i=1}^k f_i(X_i) \mid \mathcal{E}] = \prod_{i=1}^k \mathbb{E}[f_i(X_i) \mid \mathcal{E}]$$

Using induction, it suffices to show for any two bounded functions $f : \mathbb{R}^{k-1} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, we have

$$\mathbb{E}[f(X_1, \dots, X_{k-1}) g(X_k) \mid \mathcal{E}] = E[f(X_1, \dots, X_{k-1}) \mid \mathcal{E}] \mathbb{E}[g(X_k) \mid \mathcal{E}]$$

For a bounded function $\phi : \mathbb{R}^k \to \mathbb{R}$, we can define the random average

$$A_{n,k}(\phi) \triangleq \frac{1}{|I_{n,k}|} \sum_{i \in I_{n,k}} \phi(X_{i_1}, X_{i_2}, \dots, X_{i_k}).$$

It is clear that the random variable $A_{n,k}(\phi)$ is $\mathcal{E}_n$ measurable and hence $\mathbb{E}[A_{n,k}(\phi) \mid \mathcal{E}_n] = A_{n,k}(\phi)$. For each $i \in I_{n,k}$, we can find a finite permutation on $[n]$, such that $\alpha(i_j) = j$ for $j \in [k]$. Since X is exchangeable, the distribution of $(X_{i_1}, \dots, X_{i_k})$ and $(X_1, \dots, X_k)$ are identical for each $i \in I_{n,k}$. Therefore, we have

$$A_{n,k}(\phi) = \mathbb{E}[A_{n,k}(\phi) \mid \mathcal{E}_n] = \frac{1}{|I_{n,k}|} \sum_{i \in I_{n,k}} \mathbb{E}[\phi(X_{i_1}, X_{i_2}, \dots, X_{i_k}) | \mathcal{E}_n] = \mathbb{E}[\phi(X_1, X_2, \dots, X_k) | \mathcal{E}_n].$$

Since $\mathcal{E}_n \to \mathcal{E}$, using bounded convergence theorem for conditional expectations, we have

$$\lim_{n \in \mathbb{N}} A_{n,k}(\phi) = \lim_{n \in \mathbb{N}} \mathbb{E}[\phi(X_1, X_2, \dots, X_k) | \mathcal{E}_n] = \mathbb{E}[\phi(X_1, X_2, \dots, X_k) | \mathcal{E}].$$

Let f and g be real-valued bounded functions on $\mathbb{R}^{k-1}$ and $\mathbb{R}$ respectively, such that $\phi(x_1, \dots, x_k) \triangleq f(x_1, \dots, x_{k-1}) g(x_k)$. We also define $\phi_j(x_1, \dots, x_{k-1}) \triangleq f(x_1, \dots, x_{k-1}) g(x_j)$ for each $j \in [k-1]$, to write

$$\begin{aligned}
(n)_{k-1} A_{n,k-1}(f) n A_{n,1}(g) &= \sum_{i \in I_{n,k-1}} f(X_{i_1}, \dots, X_{i_{k-1}}) \sum_{m=1}^{n} g(X_m) \\
&= \sum_{i \in I_{n,k}} f(X_{i_1}, \dots, X_{i_{k-1}}) g(X_{i_k}) + \sum_{i \in I_{n,k-1}} \sum_{j=1}^{k-1} f(X_{i_1}, \dots, X_{i_{k-1}}) g(X_{i_j}) \\
&= (n)_k A_{n,k}(\phi) + \sum_{j=1}^{k-1} (n)_{k-1} A_{n,k-1}(\phi_j).
\end{aligned}$$

Dividing both sides by $(n)_k$ and rearranging terms, we get

$$A_{n,k}(\phi) = \frac{n}{n-k+1} A_{n,k-1}(f) A_{n,1}(g) - \frac{1}{n-k+1} \sum_{j=1}^{k-1} A_{n,k-1}(\phi_j),$$

Taking limits on both sides, we obtain the result

$$\mathbb{E}[f(X_1, \dots, X_{k-1}) g(X_k) | \mathcal{E}] = \mathbb{E}[f(X_1, \dots, X_{k-1}) | \mathcal{E}] \mathbb{E}[g(X_k) | \mathcal{E}].$$

□

Corollary 1.16 (De Finetti 1931). *For any $n \in \mathbb{N}$, A random binary sequence $X : \Omega \to \{0,1\}^{\mathbb{N}}$ is exchangeable iff there exists a distribution function $F : [0,1] \to [0,1]$ such that for any $n \in \mathbb{N}$, $x \in \{0,1\}^n$, and $s_n \triangleq \sum_i x_i$*

$$P(\cap_{i=1}^n \{X_i = x_i\}) = \int_0^1 p^{s_n} (1-p)^{n-s_n} dF(p).$$

Proof. For each $n \in \mathbb{N}$, we define $Y_n \triangleq \frac{1}{n} \sum_{i=1}^n X_i$. We observe that the random sequence $Y : \Omega \to [0,1]^{\mathbb{N}}$ is bounded as

$$0 \leqslant \inf_{n \in \mathbb{N}} Y_n \leqslant \sup_{n \in \mathbb{N}} Y_n \leqslant 1.$$

Hence, $0 \leqslant \liminf_n Y_n \leqslant \limsup_n Y_n \leqslant 1$ exists and are bounded. We define the limit $Y_\infty \triangleq \lim_{n \in \mathbb{N}} \frac{S_n}{n}$ when it exists. Since X is exchangeable, $\mathcal{E}_n \triangleq \sigma(X_1, \dots, X_n)$ is n-exchangeable, and thus $\mathcal{E}_n = \sigma(Y_n)$. Therefore, the exchangeable σ-algebra $\mathcal{E} = \sigma(Y_\infty)$. Let F be the distribution function for the random variable Y_∞. From Theorem 1.15, random vector $(X_1, \dots, X_n)$ is conditionally *i.i.d.* given $\sigma(Y_\infty)$, with $P(\{X_i = 1\} \mid \sigma(Y_\infty)) = Y_\infty$. Then, we can write

$$P(\cap_{i=1}^n \{X_i = x_i\}) = \mathbb{E}[\mathbb{E}[\prod_{i=1}^n \mathbb{1}_{\{X_i = x_i\}} \mid \mathcal{E}]] = \mathbb{E}[Y_\infty^{s_n} (1 - Y_\infty)^{n-s_n}] = \int_0^1 dF(p) p^{s_n} (1-p)^{n-s_n}.$$

□

Example 1.17 (Polya's Urn Scheme). We now discuss a non-trivial example of exchangeable random variables. Consider a discrete time stochastic process $\{(B_n, W_n) : n \in \mathbb{N}\}$, where B_n, W_n respectively denote the number of black and white balls in an urn after $n \in \mathbb{N}$ draws. At each draw n, balls are uniformly sampled from this urn. After each draw, one additional ball of the same color to the drawn ball, is returned to the urn. We are interested in characterizing evolution of this urn, given initial urn content (B_0, W_0). Let ξ_i be a random variable indicating the outcome of the ith draw being a black ball. For example, if the first drawn ball is a black, then $\xi_1 = 1$ and $(B_1, W_1) = (B_0 + 1, W_0)$. In general,

$$B_n = B_0 + \sum_{i=1}^{n} \xi_i = B_{n-1} + \xi_n, \qquad W_n = W_0 + \sum_{i=1}^{n} (1 - \xi_i) = W_{n-1} + 1 - \xi_n.$$

It is clear that $B_n + W_n = B_0 + W_0 + n$. Consider a random sequence of indicators $\xi : \Omega \to \{0,1\}^{[n]}$. We can find the indices of black balls being drawn in first n draws, as

$$I_n(\xi) \triangleq \{i \in [n] : \xi_i = 1\}.$$

With this, we can write the probability of the event $\{\xi = x\}$ for some binary sequence $x \in \{0,1\}^n$ as

$$P\Big(\cap_{i=1}^{n} \{\xi_i = x_i\} \Big) = \frac{\prod_{i=1}^{|I_n(x)|}(B_0 + i - 1) \prod_{j=1}^{n-|I_n(x)|}(W_0 + j - 1)}{\prod_{i=1}^{n}(B_0 + W_0 + i - 1)}$$

Since this probability depends only on $|I_n(x)|$ and not x for any $n \in \mathbb{N}$, it shows that any finite number of draws is finitely permutable event. That is, $(\xi_1, \dots, \xi_n) \in \mathcal{E}_n$ for each $n \in \mathbb{N}$. Hence, any sequence of draws ξ for Polya's Urn scheme is exchangeable.