Lecture-27: Exchangeability

1 Exchangeability

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

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

Definition 1.2. We define a projection operator $\pi_n : \prod_{i\in\mathbb{N}} \Omega_i \to \Omega_n$ such that $\pi_n(\omega) \triangleq \omega_n$ 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 $P \circ \pi_i^{-1}(A_i) = \mu_i(A_i)$ 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 $\sigma : \mathbb{N} \to \mathbb{N}$, then we can define a finitely permuted outcome $\sigma(\omega) \triangleq (\omega_{\sigma(i)} : i \in \mathbb{N})$ in terms of its projections, as

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

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

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

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.

Definition 1.8. 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 the **exchangeable** σ-algebra and denoted by $\mathcal{E}$.

Definition 1.9. A random sequence $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ is called **exchangeable** if for each n-permutation $\sigma : [n] \to [n]$, the joint distribution of $(X_1, X_2, \dots, X_n)$ and $(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)})$ 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 1.10 (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 $X : \Omega \to \mathbb{R}^{\mathbb{N}}$. Then the events $\{S_n \leqslant x \text{ infinitely often }\}$ and $\left\{\limsup_{n\in\mathbb{N}} \frac{S_n}{c_n} \geqslant 1\right\}$ are permutable for any constant sequence $c \in (\mathbb{R} \setminus \{0\})^{\mathbb{N}}$. This is due to the fact that $S_n(X) = S_n(\pi(X))$ is n-permutable.

Further, all events in tail σ-algebra $\mathcal{T}$ are permutable. This follows from the fact that if the event $A \in \sigma(X_{n+1}, X_{n+2}, \dots)$, then the event A remains unaffected by the permutation of $(X_1, \dots, X_n)$ and hence A is n- permutable.

Example 1.11 (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 $W \triangleq \{i \in [n] : \xi_i = 1\}$, where $|W| = k$. Then, 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

$$P\{W = A\} = P\{\xi_i = 1, i \in A, \xi_j = 0, j \notin A\} = \frac{k(k-1)\dots 1 \times (n-k)(n-k-1)\dots 1}{n(n-1)(n-2)\dots 1} = \frac{(n-k)!k!}{n!} = \frac{1}{\binom{n}{k}}.$$

This joint distribution is independent of set of exact locations A, and hence exchangeable.

However, one can see the dependence from

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

Example 1.12 (Conditionally independent sequence). Let $Y : \Omega \to \mathcal{Y}$ denote a discrete random variable with probability mass function $p : \mathcal{Y} \to [0,1]$. Let $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ be a conditionally *i.i.d.* random sequence given random variable Y, with conditional distribution F_y given $Y = y$. We can write the joint finite dimensional distribution of the sequence X,

$$P\{X_1 \leqslant x_1 \dots, X_n \leqslant x_n\} = \sum_{y\in\mathcal{Y}} P(\{X_1 \leqslant x_1 \dots, X_n \leqslant x_n\} \mid \{Y = y\}) P\{Y = y\} = \sum_{y\in\mathcal{Y}} \prod_{i=1}^{n} F_y(x_i) p(y).$$

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

Theorem 1.13 (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 $\mathcal{E}$.*

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

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

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

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

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

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

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

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

It is clear that the random variable $A_n(\phi)$ is $\mathcal{E}_n$ measurable and hence $\mathbb{E}[A_n(\phi)|\mathcal{E}_n] = A_n(\phi)$. For each $i \in I_{n,k}$, we can find a finite permutation on $[n]$, such that $\sigma(i_j) = j$ 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

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

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

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

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

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

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

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

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

$$P\{X_1 = x_1, \dots, X_n = x_n\} = \int_0^1 p^{s_n}(1-p)^{n-s_n} dF(p).$$

Proof. The distribution of an n-exchangeable binary sequence is given for some pmf $p : \left\{0, \frac{1}{n}, \dots, 1\right\} \to [0,1]$ by $\sum_{i=0}^{n} p_i \mathbb{1}_{\left\{\frac{S_n}{n} = \frac{i}{n}\right\}}$ where $p_i = P(S_n/n = i/n)$. Let $Y_n \triangleq \frac{S_n}{n}$ for each $n \in \mathbb{N}$ and $Y \triangleq \lim_{n \in \mathbb{N}} \frac{S_n}{n}$. Hence, the collection of n-permutable events is $\mathcal{E}_n = \sigma(Y_n)$ for binary sequences. Therefore, the exchangeable σ-algebra $\mathcal{E} = \sigma(Y)$, and F is the distribution function for the random variable Y.

Since $(X_1, \dots, X_n)$ are conditionally *i.i.d.* given $\sigma(Y)$, with $P(\{X_i = 1\}|\sigma(Y)) = Y$. Then, we can write

$$P\{X_1 = x_1, \dots, X_n = x_n\} = \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{X_1 = x_1, \dots, X_n = x_n\}}|\mathcal{E}]] = \int_{p=0}^{1} dF(p)p^{s_n}(1-p)^{n-s_n}.$$

□

Example 1.15 (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

$$P\{\xi_1 = x_1, \dots, \xi_n = x_n\} = \frac{\prod_{i=1}^{|I_n(x)|}(B_0 + i - 1)\prod_{j=1}^{n-|I_n(x)|}(W_0 + j - 1)}{\prod_{i=1}^{n}(B_0 + W_0 + i - 1)}$$

Since this probability depends only on $|I_n(x)|$ and not x, 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.