Lecture 25: Exchangeability

1 Exchangeability

Let X_i be a random variable on the probability space $(S_i, \mathcal{S}_i, \mu_i)$. Consider the probability space $(\Omega, \mathcal{F}, P)$ for the process $X = \{X_i : i \in \mathbb{N}\}$ where

$$\Omega = \prod_{i \in \mathbb{N}} S_i, \qquad \mathcal{F} = \prod_{i \in \mathbb{N}} \mathcal{S}_i.$$

Let $p_i : \Omega \to S_i$ be a projection operator, such that $p_i(\omega) = \omega_i$, then for each $i \in \mathbb{N}$

$$\mu_i = P \circ p_i^{-1}.$$

A **finite permutation** of $\mathbb{N}$ is a bijective map $\pi : \mathbb{N} \to \mathbb{N}$ such that $\pi(i) \neq i$ for only finitely many i. That is, for a finite $I \subset \mathbb{N}$, we have

$$\pi(I) = I, \qquad \pi(i) = i, \ \ i \notin I.$$

Let $\omega \in \Omega$, p_i be projection operators, and π be a finite permutation, then we can define a finitely permuted outcome $\pi(\omega)$ in terms of its projections, as

$$p_i \circ \pi(\omega) = p_{\pi(i)} \circ \omega, \ \ i \in \mathbb{N}.$$

An event $A \subset \Omega$ is **permutable** if for any finite permutation π,

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

It is clear that a finite permutation π can always be defined on an interval of form $[n]$, where $n = \max A$. The collection of permutable events is a σ-field called the **exchangeable σ-field** and denoted by $\mathcal{E}$. A sequence $X = \{X_n : n \in \mathbb{N}\}$ of random variables is called **exchangeable** if for each finite permutation π on a finite set $[n]$, the joint distribution of $(X_1, X_2, \dots, X_n)$ and $(X_{\pi(1)}, X_{\pi(2)}, \dots, X_{\pi(n)})$ are identical.

Example 1.1. Suppose balls are selected randomly, without replacement, from an urn consisting of n balls of which k are white. For $i \in [n]$, let X_i be the indicator of the event that the ith selection is white. Then the finite collection $(X_1, \dots X_n)$ is exchangeable but not independent. In particular, let $A = \{i \in [n] : X_i = 1\}$. Then, we know that $|A| = k$, and we can write

$$\Pr\{X_i = 1, i \in A, X_j = 0, j \in A^c\} = \Pr\{A = (i_1, i_2, \dots, i_k)\} = \frac{(n-k)!k!}{n!} = \frac{1}{\binom{n}{k}}.$$

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

$$\Pr\{X_2 = 1 | X_1 = 1\} = \frac{k-1}{n-1} \neq \frac{k}{n-1} = \Pr\{X_2 = 1 | X_1 = 0\}.$$

Example 1.2. Let Λ denote a random variable having distribution G. Let X be a sequence of dependent random variables, where each of these random variables are conditionally *iid* with distribution F_λ given $\Lambda = \lambda$. We can write the joint finite dimensional distribution of the sequence X,

$$\Pr\{X_1 \leq x_1 \ldots, X_n \leq x_n\} = \int_\Lambda \prod_{i=1}^{n} F_\lambda(x_i) dG(\lambda).$$

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

Theorem 1.3 (De Finetti's Theorem). *If X is an exchangeable sequence of random variables then conditioned on $\mathcal{E}$, the sequence X is iid.*

Proof. To show the independence of exchangeable sequence X of random variables, conditioned on exchangeable σ-field $\mathcal{E}$, 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 that any two bounded functions $f : \mathbb{R}^{k-1} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are independent conditioned on the exchangeable σ-field. That is,

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

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

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

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

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

It is clear that $A_n(\phi) \in \mathcal{E}_n$ 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 $\pi(i_j) = j$ for $j \in [k]$, and $\pi(j) = j \in [n] \setminus I_{n,k}$. Since X is exchangeable, the distribution of $(X_{i_1}, \ldots, X_{i_k})$ and $(X_1, \ldots, X_k)$ are identical for each $i \in I_{n,k}$. Therefore, we have

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

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

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

Let f and g be bounded functions on $\mathbb{R}^{k-1}$ and $\mathbb{R}$ respectively, such that $\phi(x_1, \ldots, x_k) = f(x_1, \ldots, x_{k-1})g(x_k)$.

We also define $\phi_j(x_1,\dots,x_{k-1}) = f(x_1,\dots,x_{k-1})g(x_j)$, to write

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

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

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

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 1.4 (De Finetti 1931). *A binary sequence $X = \{X_n : n \in \mathbb{N}\}$ is exchangeable iff there exists a distribution function $F(p)$ on $[0,1]$ such that for any $n \in \mathbb{N}$, and $S_n = \sum_i x_i$*

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

Proof. Let $Y_n = \frac{S_n}{n}$ and $Y = \lim_{n\in\mathbb{N}}\frac{S_n}{n}$. It suffices to show that for binary sequences, the collection of finitely permutable events is $\mathcal{E}_n = \sigma(Y_n)$. Hence, the exchangeable σ-field $\mathcal{E} = \sigma(Y)$, and F is the distribution function for the random variable Y. □

2 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 $B_n + W_n = B_0 + W_0 + n$. Let ξ be any sequence of indicators $\xi \in \{0,1\}^{\mathbb{N}}$. We can find the indices of black balls being drawn in first n draws, as

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

With this, we can write the probability of $x \in \{0,1\}^n$

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

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, $\xi \in \mathcal{E}_n$ for each $n \in \mathbb{N}$. Hence, any sequence of draws $\xi = \{\xi_i : i \in \mathbb{N}\}$ for Polya's Urn scheme is exchangeable.

We are interested in limiting ratio of black balls. We represent the proportion of black balls after n draws by

$$X_n = \frac{B_n}{B_n + W_n} = \frac{B_n}{B_0 + W_0 + n}.$$

Let $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ be the σ-field generated by the first n indicators to black draws. It is clear that

$$\mathbb{E}[\xi_{n+1}|\mathcal{F}_n] = X_n.$$

Using this fact, we observe that $X = \{X_n : n \in \mathbb{N}\}$ is a martingale adapted to filtration $\mathcal{F} = \{\mathcal{F}_n : n \in \mathbb{N}\}$, since

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \frac{1}{B_0 + W_0 + n + 1}\mathbb{E}[B_{n+1}|\mathcal{F}_n] = \frac{B_n + X_n}{B_0 + W_0 + n + 1} = X_n.$$

For each $n \in \mathbb{N}$, we have $\mathbb{E}X_n^+ = \mathbb{E}X_n \leq 1$. From Martingale convergence theorem, it follows almost surely that

$$\lim_{n \in \mathbb{N}} X_n = X_0 = \frac{B_0}{B_0 + W_0}.$$