Exchangeability and Optional Stopping Theorem

15.1 Exchangeability and Hewitt-Savage 0-1 Law

Definition 15.1. *Let $X_1, X_2, \ldots$ be a sequence of r.v.'s. Define,*

$$\mathcal{E}_n = \sigma(f(X_1, X_2, \ldots, X_n, \ldots, X_k),\ k \geqslant n),$$

where f is symmetric under co-ordinate permutation in the first n-coordinates. The σ-field

$$\mathcal{E}_\infty = \cap_{n\geqslant 1}\mathcal{E}_n$$

is called the exchangeable σ-field.

It can be shown that $\mathcal{E}_n = \sigma(X_1^{(n)}, \ldots, X_n^{(n)}, X_{n+1}, \ldots)$, where $X_1^{(n)} \leqslant X_2^{(n)} \leqslant \cdots \leqslant X_n^{(n)}$ are the order statistic of $X_1, \cdots, X_n$.

Lemma 15.2 (Hewitt-Savage 0-1 Law). *For a sequence of i.i.d. r.v.'s and $A \in \mathcal{E}_\infty$, $\mathbb{P}(A) = 0$ or 1.*

Proof. We will show that $\mathcal{E}_\infty \perp \sigma(X_1, X_2, \ldots, X_n)$ which will imply that $A \perp A$ and thus $\mathbb{P}(A) = 0$ or 1.

Claim: If $\mathbb{E}(\mathbb{1}_A \mid \mathcal{E}_\infty) = \mathbb{P}(A)$ then $A \perp \mathcal{E}_\infty$.

For $B \in \mathcal{E}_\infty$, $\mathbb{P}(AB) = \mathbb{E}(\mathbb{1}_A\mathbb{1}_B) = \mathbb{E}(\mathbb{E}(\mathbb{1}_A \mid \mathcal{E}_\infty)\mathbb{1}_B) = \mathbb{E}(\mathbb{P}(A)\mathbb{1}_B) = \mathbb{P}(A)\,\mathbb{P}(B)$. This implies $A \perp B$.

Note that by reverse martingale, $\mathbb{E}(\mathbb{1}_A \mid \mathcal{E}_n) \xrightarrow{a.s.} \& \xrightarrow{\mathbf{L^1}} \mathbb{E}(\mathbb{1}_A \mid \mathcal{E}_\infty)$.

Claim: For $Y = \phi(X_1, \ldots, X_k)$,

$$\mathbb{E}(Y \mid \mathcal{E}_n) = \frac{1}{\binom{n}{k}} \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} \phi(X_{i_1}, \ldots, X_{i_k}).$$

Proof: $\mathbb{E}(\phi(X_1, \ldots, X_k) \mid \mathcal{E}_n) = \mathbb{E}(\phi(X_{i_1}, \ldots, X_{i_k}) \mid \mathcal{E}_n)$, for $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$.

Claim: Take $Y = \phi(X_1, \ldots, X_k)$, then $\mathbb{E}(Y \mid \mathcal{E}_\infty) \perp X_1$.

Proof: $\mathbb{E}(Y \mid \mathcal{E}_n) = \frac{1}{\binom{n}{k}}(\text{terms that depend on } X_1) + \frac{1}{\binom{n}{k}}(\text{functions of } X_2, \ldots, X_n)$. Now number of terms that depends on X_1 is $\leqslant n^{k-1}$. Moreover, we have $\mathbb{E}(Y \mid \mathcal{E}_n) \xrightarrow{a.s.} \mathbb{E}(Y \mid \mathcal{E}_\infty)$ and $|\frac{1}{\binom{n}{k}}(\text{terms that depend on } X_1)| \leqslant \frac{n^{k-1}}{\binom{n}{k}} \to 0$. Thus $\mathbb{E}(Y \mid \mathcal{E}_\infty) \in \sigma(X_2, X_3, \ldots)$. Same argument implies, for any $X_\ell, \ell \geqslant 1$, we have $\mathbb{E}(Y \mid \mathcal{E}_\infty) \perp \sigma(X_1, \ldots, X_\ell)$, and thus $\mathbb{E}(Y \mid \mathcal{E}_\infty) \perp \sigma(X_1, X_2, \ldots)$, which implies

$$\mathbb{E}(Y \mid \mathcal{E}_\infty) \perp \mathbb{E}(Y \mid \mathcal{E}_\infty), \text{ and } \mathbb{E}(Y \mid \mathcal{E}_\infty) = \mathbb{E}\,Y.$$

■

Definition 15.3. *We say that* $(X_1, X_2, \cdots)$ *a sequence of random variables is exchangeable if*

$$(X_1, \dots, X_n) \stackrel{\mathrm{d}}{=} (X_{\pi_1}, \dots, X_{\pi_n}),$$

where $\pi \in \mathfrak{S}_n = \{\tau \mid \tau : \{1, \dots, n\} \to \{1, \dots, n\} \text{ is a bijection}\}$.

Theorem 15.4 (de Finetti's Theorem)**.** *Let* $X_1, X_2, \dots$ *be an exchangeable sequence of random variables. Then*

$$\mathbb{P}(X_i \in A_i \,, i = 1, \dots, n \mid \mathcal{E}_\infty) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i \mid \mathcal{E}_\infty),$$

where $\mathcal{E}_\infty = \cap_{n \geqslant 1} \sigma(X_1^{(n)}, \dots, X_n^{(n)}, X_{n+1}, \dots)$ *is the exchangeable* σ*-field.*

Corollary 15.5. *Let* $(X_1, \cdots, X_n, \cdots)$ *be exchangeable and* $X_i \in \{0, 1\}$ $\forall i$*, then there exists a r.v.* $\Theta \in [0, 1]$ *which is* $\sim \mathcal{E}_\infty$ *measurable, s.t.*

$$\mathbb{P}(X_1 = i_1, \dots, X_k = i_k \mid \mathcal{E}_\infty) = \Theta^{\#\{i_j = 1\}}(1 - \Theta)^{n - \#\{i_j = 1\}},$$

where $\Theta = \mathbb{P}(X_1 = 1 \mid \mathcal{E}_\infty)$.

15.2 Optional Stopping Theorem

Lemma 15.6. *If* $(\mathcal{F}_n)_{n \geqslant 0}$ *is a filtration and* N *is a stopping time, then*

$$\mathcal{F}_N = \{A \in \mathcal{F} \mid A \cap \{N = i\} \in \mathcal{F}_i \text{ for all } i \geqslant 0\}$$

is a σ*-field.*

Remark 15.7. *If* $N = k$ *a.s., then* $\mathcal{F}_N = \mathcal{F}_k$*. Moreover,* $N \leqslant M$ *implies* $\mathcal{F}_N \subseteq \mathcal{F}_M$*.*

Lemma 15.8. *Let* N *be a stopping time with respect to* $(\mathcal{F}_n)$ *such that* $0 \leqslant T \leqslant N \leqslant l$ *a.s. Let* $(X_n, \mathcal{F}_n)_{n \geqslant 1}$ *be a sub-martingale. Then,*

(i) $X_N \leqslant \mathbb{E}(X_l \mid \mathcal{F}_N)$.

(ii) $\mathbb{E}(X_T) \leqslant \mathbb{E}(X_N) \leqslant \mathbb{E}(X_l)$,

Proof. Take $A \in \mathcal{F}_N$. We want to prove that $\mathbb{E}(X_l \mathbb{1}_A) \geqslant \mathbb{E}(X_N \mathbb{1}_A)$. Since $N \leqslant l$, it is enough to show that $\mathbb{E}(X_l \mathbb{1}_{A \cap \{N = i\}}) \geqslant \mathbb{E}(X_i \mathbb{1}_{A \cap \{N = i\}})$ for $i = 0, 1, \dots, l$. Now, result (i) follows by the fact that $A \cap \{N = i\} \in \mathcal{F}_i$ for all i.

Taking expectation in (i), we get $\mathbb{E}(X_N) \leqslant \mathbb{E}(X_l)$. For the other inequality, note that $X_{N \wedge n} - X_{T \wedge n} = \sum_{i=1}^{n} \Delta X_i \mathbb{1}_{T < i \leqslant N}$. Now, X_n is a sub-martingale and $\mathbb{1}_{T < i \leqslant N} = \mathbb{1}_{T \leqslant i-1 < N}, i \geqslant 1$ is predictable, hence $X_{N \wedge n} - X_{T \wedge n}$ is a sub-martingale, which implies that $\mathbb{E}(X_{N \wedge n} - X_{T \wedge n}) \geqslant 0$ and take $n = l$ to get the other inequality in (ii). ■

If we assume uniform integrability, so that $X_\infty = \lim_{n \to \infty} X_n$ exists in a.s. and L^1 sense, we can prove more.

Theorem 15.9. *If* $(X_n, \mathcal{F}_n)_{n \geqslant 0}$ *is a uniformly integrable submartingale, then for any stopping time* N*,* $(X_{N \wedge n})_{n \geqslant 0}$ *is uniformly integrable.*

Proof. The proof involves showing that $\mathbb{E}|X_N| < \infty$, since $|X_{N\wedge n}| = |X_N|\mathbb{1}_{\{N\leqslant n\}} + |X_n|\mathbb{1}_{\{N>n\}}$ combined with u.i. of X_n proves the result. We note that $\mathbb{E}\, X^+_{N\wedge n} \leqslant \mathbb{E}\, X^+_n$ for $n \geqslant 0$ to complete the proof. ■

Theorem 15.10 (Optional Stopping Theorem). *If $L \leqslant M$ are stopping times w.r.t. the filtration $(\mathcal{F}_n)_{n\geqslant 0}$ and $(X_{M\wedge n}, \mathcal{F}_n)_{n\geqslant 0}$ is a uniformly integrable submartingale, then*

$$X_L \leqslant \mathbb{E}(X_M \mid \mathcal{F}_L) \text{ and } \mathbb{E}\, X_L \leqslant \mathbb{E}\, X_M.$$

Proof. Enough to show that $\mathbb{E}((X_M - X_L)\mathbb{1}_A) \geqslant 0$ for all $A \in \mathcal{F}_L \subseteq \mathcal{F}_M$. Fix $A \in \mathcal{F}_L$ and define $N = M\mathbb{1}_{A^c} + L\mathbb{1}_A$.

We claim that N is a stopping time and $\mathbb{E}(X_M - X_N) \geqslant 0$. The first claim follows since $\{N = n\} = (A \cap \{L = n\}) \cup (\{M = n\} \cap A^c) \in \mathcal{F}_n$.

The second claim follows since $(X_{M\wedge n})_{n\geqslant 0}$ is a uniformly integrable submartingale and $\mathbb{E}(X_{M\wedge n} - X_{N\wedge n}) \geqslant 0$ for all $n \geqslant 0$. Now note that $X_M - X_N = (X_M - X_L)\mathbb{1}_A$ and we have the proof. ■

Corollary 15.11 (Generalization of Wald's First Identity). *Suppose $(X_n)_{n\geqslant 0}$ is a submartingale and $\sup_{n\geqslant 0} \mathbb{E}(|X_{n+1} - X_n| \mid \mathcal{F}_n) \leqslant B < \infty$ a.s. If N is a stopping time with $\mathbb{E}\, N < \infty$, then $(X_{N\wedge n}, n \geqslant 0)$ is uniformly integrable and hence $\mathbb{E}\, X_N \geqslant \mathbb{E}\, X_0$. Moreover, if $(X_n)_{n\geqslant 0}$ is also a martingale, then $\mathbb{E}\, X_N = \mathbb{E}\, X_0$.*

15.3 Azuma-Hoeffding Inequality

Theorem 15.12 (Azuma-Hoeffding Inequality). *Let $(M_n, \mathcal{F}_n)_{n\geqslant 1}$ be a super-martingale with the Martingale Difference Sequence $\Delta_n := M_n - M_{n-1}$ satisfying*

$$|\Delta_n| \leqslant c_n \text{ for all } n \geqslant 1.$$

Then for all $t \geqslant 0$ we have

$$\mathbb{P}(M_n - M_0 \geqslant t) \leqslant e^{-t^2/2s_n^2}$$

where

$$s_n^2 := \sum_{i=1}^{n} c_i^2, n \geqslant 1.$$

In particular, if $(M_n, \mathcal{F}_n)_{n\geqslant 1}$ is a martingale, we have

$$\mathbb{P}(|M_n - M_0| \geqslant t) \leqslant 2e^{-t^2/2s_n^2}, t \geqslant 0.$$

Proof. We will use the following result: For $0 \leqslant |x| \leqslant c$, we have

$$e^x \leqslant \frac{1}{2}(e^c + e^{-c}) + \frac{1}{2c}(e^c - e^{-c}) \cdot x \leqslant e^{c^2/2} + \frac{\sinh(c)}{c} \cdot x.$$

Note that $\mathbb{E}(\Delta_n \mid \mathcal{F}_{n-1}) \leqslant 0$ for all $n \geqslant 1$. Thus for any $\theta > 0, n \geqslant 1$, we have

$$\mathbb{E}(e^{\theta\Delta_n} \mid \mathcal{F}_{n-1}) \leqslant e^{\theta^2 c_n^2/2}.$$

Using induction, we have

$$\mathbb{E}(e^{\theta(M_n - M_0)}) \leqslant \prod_{i=1}^{n} e^{\theta^2 c_i^2/2} = e^{\theta^2 s_n^2/2}.$$

In particular, using exponential Markov's inequality we have

$$\mathbb{P}(M_n - M_0 \geqslant t) \leqslant \inf_{\theta > 0} e^{-\theta t}\, \mathbb{E}(e^{\theta(M_n - M_0)}) \leqslant \inf_{\theta > 0} e^{-\theta t + \theta^2 s_n^2/2} = e^{-t^2/2s_n^2}.$$

■