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,
\[
\mathcal{E}_n = \sigma(f(X_1, X_2, \ldots, X_n, \ldots, X_k), \ k \geq n),
\]
where \( f \) is symmetric under co-ordinate permutation in the first \( n \)-coordinates. The \( \sigma \)-field
\[
\mathcal{E}_\infty = \bigcap_{n \geq 1} \mathcal{E}_n
\]
is called the exchangeable \( \sigma \)-field.

It can be shown that \( \mathcal{E}_n = \sigma(X^{(n)}_1, \ldots, X^{(n)}_n, X_{n+1}, \ldots) \), where \( X^{(n)}_1 \leq X^{(n)}_2 \leq \cdots \leq X^{(n)}_n \) are the order statistic of \( X_1, \ldots, X_n \).

**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 \( \mathbb{E}(\mathbb{1}_A | \mathcal{E}_\infty) = \mathbb{P}(A) \) then \( A \perp \mathcal{E}_\infty \).

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

Note that by reverse martingale, \( \mathbb{E}(\mathbb{1}_A | \mathcal{E}_n) \overset{a.s.}{\longrightarrow} & \overset{L^1}{\longrightarrow} \mathbb{E}(\mathbb{1}_A | \mathcal{E}_\infty) \).

**Claim:** For \( Y = \phi(X_1, \ldots, X_k) \),
\[
\mathbb{E}(Y | \mathcal{E}_n) = \frac{1}{\binom{n}{k}} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \phi(X_{i_1}, \ldots, X_{i_k}).
\]

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

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

Proof: \( \mathbb{E}(Y | \mathcal{E}_n) = \frac{1}{\binom{n}{k}} \) (terms that depend on \( X_1 \)) + \( \frac{1}{\binom{k}{k}} \) (functions of \( X_2, \ldots, X_n \)). Now number of terms that depends on \( X_1 \) is \( \leq n^{k-1} \). Moreover, we have \( \mathbb{E}(Y | \mathcal{E}_n) \overset{a.s.}{\longrightarrow} \mathbb{E}(Y | \mathcal{E}_\infty) \) and \( | \frac{1}{\binom{k}{k}} \) (terms that depend on \( X_1 \)) \( \leq \frac{n^{k-1}}{\binom{k}{k}} \rightarrow 0 \). Thus \( \mathbb{E}(Y | \mathcal{E}_\infty) \in \sigma(X_2, X_3, \ldots) \). Same argument implies, for any \( X_\ell, \ell \geq 1 \), we have \( \mathbb{E}(Y | \mathcal{E}_\infty) \perp \sigma(X_1, X_2, \ldots) \), and thus \( \mathbb{E}(Y | \mathcal{E}_\infty) \perp \sigma(X_1, X_2, \ldots) \), which implies
\[
\mathbb{E}(Y | \mathcal{E}_\infty) \perp \mathbb{E}(Y | \mathcal{E}_\infty), \text{ and } \mathbb{E}(Y | \mathcal{E}_\infty) = \mathbb{E} Y.
\]
**Definition 15.3.** We say that \((X_1, X_2, \ldots)\) a sequence of random variables is exchangeable if
\[
(X_1, \ldots, X_n) \overset{d}{=} (X_{\pi_1}, \ldots, X_{\pi_n}),
\]
where \(\pi \in \mathfrak{S}_n = \{\tau \mid \tau : \{1, \ldots, n\} \to \{1, \ldots, n\} \text{ is a bijection}\}.

**Theorem 15.4** (de Finetti’s Theorem). Let \((X_1, X_2, \ldots)\) be an exchangeable sequence of random variables. Then
\[
\mathbb{P}(X_i \in A_i, i = 1, \ldots, n \mid \mathcal{E}_\infty) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i \mid \mathcal{E}_\infty),
\]
where \(\mathcal{E}_\infty = \bigcap_{n \geq 1} \sigma(X_1^{(n)}, \ldots, X_n^{(n)}, X_{n+1}, \ldots)\) is the exchangeable \(\sigma\)-field.

**Corollary 15.5.** Let \((X_1, \ldots, X_n, \ldots)\) 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.
\[
\mathbb{P}(X_1 = i_1, \ldots, X_k = i_k \mid \mathcal{E}_\infty) = \Theta^\#(i_1=1)(1-\Theta)^{n-\#(i_1=1)},
\]
where \(\Theta = \mathbb{P}(X_1 = 1 \mid \mathcal{E}_\infty)\).

### 15.2 Optional Stopping Theorem

**Lemma 15.6.** If \((\mathcal{F}_n)_{n \geq 0}\) is a filtration and \(N\) is a stopping time, then
\[
\mathcal{F}_N = \{A \in \mathcal{F} \mid A \cap \{N = i\} \in \mathcal{F}_i \text{ for all } i \geq 0\}
\]
is a \(\sigma\)-field.

**Remark 15.7.** If \(N = k\) a.s., then \(\mathcal{F}_N = \mathcal{F}_k\). Moreover, \(N \leq M\) 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 \(0 \leq T \leq N \leq l\) a.s. Let \((X_n, \mathcal{F}_n)_{n \geq 1}\) be a sub-martingale. Then,

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

(ii) \(\mathbb{E}(X_T) \leq \mathbb{E}(X_N) \leq \mathbb{E}(X_l)\).

**Proof.** Take \(A \in \mathcal{F}_N\). We want to prove that \(\mathbb{E}(X_l 1_A) \geq \mathbb{E}(X_N 1_A)\). Since \(N \leq l\), it is enough to show that \(\mathbb{E}(X_l 1_{A \cap \{N=i\}}) \geq \mathbb{E}(X_N 1_{A \cap \{N=i\}})\) for \(i = 0, 1, \ldots, l\). 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 \(\mathbb{E}(X_N) \leq \mathbb{E}(X_l)\). For the other inequality, note that \(X_N - X_T\mid_{T \wedge n} = \sum_{i=1}^n \Delta X_i 1_{T \leq i \leq N}\). Now, \(X_n\) is a sub-martingale and \(1_{T \leq i \leq N} = 1_{T \leq i \leq N - 1}, i \geq 1\) is predictable, hence \(X_N - X_T\mid_{T \wedge n}\) is a sub-martingale, which implies that \(\mathbb{E}(X_N - X_T\mid_{T \wedge n}) \geq 0\) and take \(n = l\) to get the other inequality in (ii). \(\blacksquare\)

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 \((X_n, \mathcal{F}_n)_{n \geq 0}\) is a uniformly integrable submartingale, then for any stopping time \(N\), \((X_N \mid_{T \wedge n})_{n \geq 0}\) is uniformly integrable.
Proof. The proof involves showing that \( E|X_N| < \infty \), since \( |X_{N\wedge n}| = |X_N|\mathbb{1}_{\{N \leq n\}} + |X_n|\mathbb{1}_{\{N > n\}} \) combined with u.i. of \( X_n \) proves the result. We note that \( E X_{N\wedge n}^+ \leq E X_n^+ \) for \( n \geq 0 \) to complete the proof. \( \blacksquare \)

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

\[
X_L \leq E(X_M | \mathcal{F}_L) \text{ and } E X_L \leq E X_M.
\]

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

We claim that \( N \) is a stopping time and \( E(X_M - X_N) \geq 0 \). 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 \( (X_{M \wedge n})_{n \geq 0} \) is a uniformly integrable submartingale and \( E(X_{M \wedge n} - X_{N \wedge n}) \geq 0 \) for all \( n \geq 0 \). Now note that \( X_M - X_N = (X_M - X_L)\mathbb{1}_A \) and we have the proof. \( \blacksquare \)

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

### 15.3 Azuma-Hoeffding Inequality

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

\[
|\Delta_n| \leq c_n \text{ for all } n \geq 1.
\]

Then for all \( t \geq 0 \) we have

\[
P(M_n - M_0 \geq t) \leq e^{-t^2/2s_n^2}
\]

where

\[
s_n^2 := \sum_{i=1}^{n} c_i^2, n \geq 1.
\]

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

\[
P(|M_n - M_0| \geq t) \leq 2e^{-t^2/2s_n^2}, t \geq 0.
\]

Proof. We will use the following result: For \( 0 \leq |x| \leq c \), we have

\[
e^x \leq \frac{1}{2}(e^c + e^{-c}) + \frac{1}{2c}(e^c - e^{-c}) \cdot x \leq e^{c^2/2} + \frac{\sinh(c)}{c} \cdot x.
\]

Note that \( E(\Delta_n | \mathcal{F}_{n-1}) \leq 0 \) for all \( n \geq 1 \). Thus for any \( \theta > 0, n \geq 1 \), we have

\[
E(e^{\theta \Delta_n} | \mathcal{F}_{n-1}) \leq e^{\theta^2 c^2_n/2}.
\]
Using induction, we have

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

In particular, using exponential Markov’s inequality we have

$$\mathbb{P}(M_n - M_0 \geq t) \leq \inf_{\theta > 0} e^{-\theta t} e^{\theta^2 s_n^2 / 2} = e^{-t^2 / 2s_n^2}.$$ 

$\blacksquare$