Section 13


Let \((\Omega, \mathcal{B}, \mathbb{P})\) be a probability space and let \((T, \leq)\) be a linearly ordered set. Consider a family of \(\sigma\)-algebras \(\mathcal{B}_t, t \in T\) such that for \(t \leq u\), \(\mathcal{B}_t \subseteq \mathcal{B}_u \subseteq \mathcal{B}\).

**Definition.** A family \((X_t, \mathcal{B}_t)_{t \in T}\) is called a martingale if

1. \(X_t : \Omega \rightarrow \mathbb{R}\) is measurable w.r.t. \(\mathcal{B}_t\); in other words, \(X_t\) is adapted to \(\mathcal{B}_t\).
2. \(\mathbb{E}|X_t| < \infty\).
3. \(\mathbb{E}(X_u|\mathcal{B}_t) = X_t\) for \(t \leq u\).

If the last equality is replaced by \(\mathbb{E}(X_u|\mathcal{B}_t) \leq X_t\) then the process is called a supermartingale and if \(\mathbb{E}(X_u|\mathcal{B}_t) \geq X_t\) then it is called a submartingale.

**Examples.**

1. Consider a sequence \((X_n)_{n \geq 1}\) of independent random variables such that \(\mathbb{E}X_1 = 0\) and let \(S_n = \sum_{i \leq n} X_i\). If \(\mathcal{B}_n = \sigma(X_1, \ldots, X_n)\) is a \(\sigma\)-algebra generated by the first \(n\) r.v.s then \((S_n, \mathcal{B}_n)_{n \geq 1}\) is a martingale since
   \[
   \mathbb{E}(S_{n+1}|\mathcal{B}_n) = \mathbb{E}(X_{n+1} + S_n|\mathcal{B}_n) = 0 + S_n = S_n.
   \]
2. Consider a sequence of \(\sigma\)-algebras
   \[
   \ldots \subseteq \mathcal{B}_m \subseteq \mathcal{B}_n \subseteq \ldots \subseteq \mathcal{B}
   \]
   and a r.v. \(X\) on \(\mathcal{B}\) and let \(X_n = \mathbb{E}(X|\mathcal{B}_n)\). Then \((X_n, \mathcal{B}_n)\) is a martingale since for \(m < n\)
   \[
   \mathbb{E}(X_n|\mathcal{B}_m) = \mathbb{E}(\mathbb{E}(X|\mathcal{B}_n)|\mathcal{B}_m) = \mathbb{E}(X|\mathcal{B}_m) = X_m.
   \]

**Definition.** If \((X_n, \mathcal{B}_n)\) is a martingale and for some r.v. \(X, X_n = \mathbb{E}(X|\mathcal{B}_n)\), then the martingale is called right-closable. If \(X_\infty = X, \mathcal{B}_\infty = \mathcal{B}\) then \((X_n, \mathcal{B}_n)_{n \leq \infty}\) is called right-closed.

3. Let \((X_i)_{i \geq 1}\) be i.i.d. and let \(S_n = \sum_{i \leq n} X_i\). Let us take \(T = \{\ldots, -2, -1\}\) and for \(n \geq 1\) define
   \[
   \mathcal{B}_- = \sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n, X_{n+1}, X_{n+2}, \ldots).
   \]
   Clearly, \(\mathcal{B}_{-(n+1)} \subseteq \mathcal{B}_-\). For \(1 \leq k \leq n\), by symmetry,
   \[
   \mathbb{E}(X_1|\mathcal{B}_-) = \mathbb{E}(X_k|\mathcal{B}_-).
   \]
   Therefore,
   \[
   S_n = \mathbb{E}(S_n|\mathcal{B}_-) = \sum_{1 \leq k \leq n} \mathbb{E}(X_k|\mathcal{B}_-) = n\mathbb{E}(X_1|\mathcal{B}_-) \implies Z_n := \frac{S_n}{n} = \mathbb{E}(X_1|\mathcal{B}_-).
   \]
Thus, \((Z_{-n}, \mathcal{B}_{-n})_{-n \leq -1}\) is a right-closed martingale.

\[\square\]

**Lemma 28** Let \(f : \mathbb{R} \to \mathbb{R}\) be a convex function. Suppose that either one of two conditions holds:

1. \((X_t, \mathcal{B}_t)\) is a martingale,
2. \((X_t, \mathcal{B}_t)\) is a submartingale and \(f\) is increasing.

Then \((f(X_t), \mathcal{B}_t)\) is a submartingale.

**Proof.**

1. For \(t \leq u\), by Jensen’s inequality,
   \[
   f(X_t) = f(E(X_u|\mathcal{B}_t)) \leq E(f(X_u)|\mathcal{B}_t).
   \]
2. For \(t \leq u\), since \(X_t \leq E(X_u|\mathcal{B}_t)\) and \(f\) is increasing,
   \[
   f(X_t) \leq f(E(X_u|\mathcal{B}_t)) \leq E(f(X_u)|\mathcal{B}_t),
   \]
where the last step is again Jensen’s inequality.

\[\square\]

**Theorem 30** *(Doob’s decomposition)* If \((X_n, \mathcal{B}_n)_{n \geq 0}\) is a submartingale then it can be uniquely decomposed \(X_n = Z_n + Y_n\),

where \((Y_n, \mathcal{B}_n)\) is a martingale, \(Z_0 = 0, Z_n \leq Z_{n+1}\) almost surely and \(Z_n\) is \(\mathcal{B}_{n-1}\)-measurable.

**Proof.**

Let \(D_n = X_n - X_{n-1}\) and

\[
G_n = E(D_n|\mathcal{B}_{n-1}) = E(X_n|\mathcal{B}_{n-1}) - X_{n-1} \geq 0
\]

by the definition of submartingale. Let,

\[
H_n = D_n - G_n, \quad Y_n = H_1 + \ldots + H_n, \quad Z_n = G_1 + \ldots + G_n.
\]

Since \(G_n \geq 0\) a.s., \(Z_n \leq Z_{n+1}\) and, by construction, \(Z_n\) is \(\mathcal{B}_{n-1}\)-measurable. We have,

\[
E(H_n|\mathcal{B}_{n-1}) = E(D_n|\mathcal{B}_{n-1}) - G_n = 0
\]

and, therefore, \(E(Y_n|\mathcal{B}_{n-1}) = Y_{n-1}\). Uniqueness follows by construction. Suppose that \(X_n = Z_n + Y_n\) with all stated properties. First, since \(Z_0 = 0, Y_0 = X_0\). By induction, given a unique decomposition up to \(n-1\),

we can write

\[
Z_n = E(Z_n|\mathcal{B}_{n-1}) = E(X_n - Y_n|\mathcal{B}_{n-1}) = E(X_n|\mathcal{B}_{n-1}) - Y_{n-1}
\]

and \(Y_n = X_n - Z_n\).

\[\square\]

**Definition.** We say that \((X_n)_{n \geq 1}\) is uniformly integrable if

\[
\sup_n E|X_n| < \infty \quad \text{and} \quad \sup_n E|X_n|I(|X_n| > M) \to 0 \quad \text{as} \quad M \to \infty.
\]

**Lemma 29** The following holds.

1. If \((X_n, \mathcal{B}_n)\) is a right-closable martingale then \((X_n)\) is uniformly integrable.
2. If \((X_n, \mathcal{B}_n)_{n \leq \infty}\) is a submartingale then for any \(a \in \mathbb{R}\), \((\max(X_n, a))\) is uniformly integrable.
Proof. 1. If $X_n = \mathbb{E}(Y|\mathcal{B}_n)$ then

$$|X_n| = |\mathbb{E}(Y|\mathcal{B}_n)| \leq \mathbb{E}(|Y|\mathcal{B}_n) \quad \text{and} \quad \mathbb{E}|X_n| \leq \mathbb{E}|Y| < \infty.$$ 

Since $\{|X_n| > M\} \in \mathcal{B}_n$,

$$X_nI(|X_n| > M) = I(|X_n| > M)\mathbb{E}(Y|\mathcal{B}_n) = \mathbb{E}(Y1(|X_n| > M)|\mathcal{B}_n)$$

and, therefore,

$$\mathbb{E}|X_n|I(|X_n| > M) \leq \mathbb{E}Y|I(|X_n| > M) \leq K\mathbb{P}(|X_n| > M) + \mathbb{E}|Y|I(|Y| > K)$$

$$\leq K\mathbb{E}|X_n|/M + \mathbb{E}|Y|I(|Y| > K) \leq K\mathbb{E}|Y|/M + \mathbb{E}|Y|I(|Y| > K).$$

Letting $M \to \infty$, $K \to \infty$ proves that $\sup_n \mathbb{E}|X_n|I(|X_n| > M) \to 0$ as $M \to \infty$.

2. Since $(X_n, \mathcal{B}_n)_{n < \infty}$ is a submartingale, for $Y = X_\infty$ we have $X_n \leq \mathbb{E}(Y|\mathcal{B}_n)$. Below we will use the following observation. Since a function $\max(a, x)$ is convex and increasing in $x$, by Jensen’s inequality

$$\max(a, X_n) \leq \mathbb{E}(\max(a, Y)|\mathcal{B}_n). \quad (13.0.1)$$

Since,

$$|\max(X_n, a)| \leq |a| + X_nI(X_n > |a|)$$

and $\{|X_n| > |a|\} \in \mathcal{B}_n$ we can write

$$\mathbb{E}|\max(X_n, a)| \leq |a| + \mathbb{E}X_nI(X_n > |a|) \leq |a| + \mathbb{E}Y1(X_n > |a|) \leq |a| + \mathbb{E}|Y| < \infty.$$

If we take $M > |a|$ then

$$\mathbb{E}|\max(X_n, a)|I(|\max(X_n, a)| > M) = \mathbb{E}X_nI(X_n > M) \leq \mathbb{E}Y1(X_n > M)$$

$$\leq K\mathbb{P}(X_n > M) + \mathbb{E}|Y|I(|Y| > K)$$

$$\leq K\mathbb{E}\max(X_n, 0) + \mathbb{E}|Y|I(|Y| > K)$$

by (13.0.1) \quad \leq K\frac{\mathbb{E}\max(Y, 0)}{M} + \mathbb{E}|Y|I(|Y| > K).$$

Letting $M \to \infty$ and $K \to \infty$ finishes the proof.

Uniform integrability plays an important role when studying the convergence of martingales. The following strengthening of the dominated convergence theorem will be useful.

Lemma 30 Consider r.v.s $(X_n)$ and $X$ such that $\mathbb{E}|X_n| < \infty, \mathbb{E}|X| < \infty$. Then the following are equivalent:

1. $\mathbb{E}|X_n - X| \to 0$,

2. $(X_n)$ is uniformly integrable and $X_n \to X$ in probability.

Proof. 2$\Rightarrow$1. We can write,

$$\mathbb{E}|X_n - X| \leq \varepsilon + \mathbb{E}|X_n - X||I(|X_n - X| > \varepsilon)$$

$$\leq \varepsilon + 2K\mathbb{P}(|X_n - X| > \varepsilon) + 2\mathbb{E}|X_n||I(|X_n| > K) + 2\mathbb{E}|X||I(|X| > K)$$

$$\leq \varepsilon + 2K\mathbb{P}(|X_n - X| > \varepsilon) + 2\sup_n \mathbb{E}|X_n||I(|X_n| > K) + 2\mathbb{E}|X||I(|X| > K).$$

Letting $n \to \infty$ and then $\varepsilon \to 0, K \to \infty$ proves the result.

1$\Rightarrow$2. By Chebyshev’s inequality,

$$\mathbb{P}(|X_n - X| > \varepsilon) \leq \frac{1}{\varepsilon} \mathbb{E}|X_n - X| \to 0$$
as \( n \to \infty \) so \( X_n \to X \) in probability. To prove uniform integrability let us first show that for any \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that

\[
P(A) < \delta \implies E[X|I_A] < \varepsilon.
\]

Suppose not. Then, for some \( \varepsilon > 0 \) one can find a sequence of events \( A(n) \) such that

\[
P(A(n)) \leq \frac{1}{2^n} \quad \text{and} \quad E[X|I_{A(n)}] > \varepsilon.
\]

Since \( \sum_{n \geq 1} P(A(n)) < \infty \), by Borel-Cantelli lemma, \( P(A(n) \text{ i.o.}) = 0 \). This means that \( |X|I_{A(n)} \to 0 \) almost surely and by the dominated convergence theorem \( E[X|I_{A(n)}] \to 0 \) - a contradiction.

Given \( \varepsilon > 0 \), take \( \delta \) as above and take \( M > 0 \) large enough so that for all \( n \geq 1 \)

\[
P(|X_n| > M) \leq \frac{E|X_n|}{M} < \delta.
\]

Then,

\[
E|X_n|I(|X_n| > M) \leq E|X_n - X| + E|X|I(|X_n| > M) \leq E|X_n - X| + \varepsilon.
\]

For large enough \( n \geq n_0 \), \( E|X_n - X| \leq \varepsilon \) and, therefore,

\[
E|X_n|I(|X_n| > M) \leq 2\varepsilon.
\]

We can also choose \( M \) large enough so that \( E|X_n|I(|X_n| > M) \leq 2\varepsilon \) for \( n \leq n_0 \) and this finishes the proof. \( \square \)