1 Conditional Expectation

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Suppose $X : \Omega \to \mathbb{R}$ is integrable and $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ-field.

Definition 1.1 *A r.v. $Y : \Omega \to \mathbb{R}$ is called the conditional expectation of X given $\mathcal{G}$, denoted by $E[X \mid \mathcal{G}]$ iff*

1. *Y is $\mathcal{G}$ measurable and integrable*
2. *$\forall\ G \in \mathcal{G}$, $\int\limits_G Y\ dP = \int\limits_G X\ dP$.*

Theorem 1.2 *If X is integrable then $E[X \mid \mathcal{G}]$ exists. If Y' is any other random variable satisfying 1) and 2) of the definition, then $Y' = E[X \mid \mathcal{G}]$ a.s. P.*

Proof: Let $X \geq 0$. Let $\gamma(G) = \int\limits_G X\ dP$. Then $\gamma \ll P$ on $\mathcal{G}$. Take $E[X \mid \mathcal{G}] := \frac{d\gamma}{dP}$. For general X take $E[X \mid \mathcal{G}] := E[X^+ \mid \mathcal{G}] - E[X^- \mid \mathcal{G}]$. Uniqueness follows from the uniqueness of the Radon-Nikodym theorem.

Remark 1.3

1. *If $\mathcal{G} = \{\phi, \Omega\}$, $E[X \mid \mathcal{G}] = EX$. If $\mathcal{G} = \mathcal{F}$, $E[X \mid \mathcal{G}] = X$ a.s.*
2. *Suppose $\mathcal{G} = \sigma\{B_i, i \geq 1\}$, $B_i \cap B_j = \phi$, $i \neq j$, $\Omega = \bigcup\limits_i B_i$*

$$\begin{aligned} E[X \mid \mathcal{G}](\omega) &:= \frac{1}{P(B_i)} \int\limits_{B_i} X\ dP,\ \omega \in B_i,\ P(B_i) > 0 \\ &:= c \qquad \textit{if } P(B_i) = 0,\ c \in \mathbb{R},\ \textit{a constant.} \end{aligned}$$

3. *If $X = I_A$, $A \in \mathcal{F}$, then $E[X \mid \mathcal{G}] = E[I_A \mid \mathcal{G}] =: P(A \mid \mathcal{G})$ a.s.*

Properties: Let X, Y and X_n be integrable.

1. If $c \in \mathbb{R}$ and $X = c$ a.s. then $E[X \mid \mathcal{G}] = c$ a.s.
2. $E[aX + bY \mid \mathcal{G}] = aE[X \mid \mathcal{G}] + bE[Y \mid \mathcal{G}]$ a.s. $a, b \in \mathbb{R}$.
3. If $X \leq Y$ a.s., then $E[X \mid \mathcal{G}] \leq E[Y \mid \mathcal{G}]$ a.s.
4. $|E[X \mid \mathcal{G}]| \leq E[|X| \mid \mathcal{G}]$ a.s.

5. If $X_n \to X$ a.s., $|X_n| \le |Y|$ a.s. then $E[X_n \mid \mathcal{G}] \to E[X \mid \mathcal{G}]$ a.s.

6. If X is $\mathcal{G}$-measurable and Y and XY are integrable then

$$E[XY \mid \mathcal{G}] = XE[Y \mid \mathcal{G}].$$

Proof: The property is first verified when $X = I_E,\ E \in \mathcal{G}$. By linearity of the conditional expectation it is then verified when X is a simple function. In general, $\exists\ X_n$, simple, $\mathcal{G}$ measurable,

$$|X_n| \le |X|,\ X_n \to X \text{ a.s. } \Rightarrow |X_n Y| \le |XY|$$

and $X_n Y \to XY$ a.s. By 5), $E[X_n Y \mid \mathcal{G}] \to E[XY \mid \mathcal{G}]$ a.s. But $E[X_n Y \mid \mathcal{G}] = X_n E[Y \mid \mathcal{G}] \to XE[Y \mid \mathcal{G}]$ a.s.

7. If $\mathcal{G}_1 \subset \mathcal{G}_2$ then $E[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1] = E[X \mid \mathcal{G}_1]$ a.s.

Proof: $\forall\ G \in \mathcal{G}_1$,

$$\int_G E[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1]\ dP = \int_G E[X \mid \mathcal{G}_2]\ dP = \int_G X\ dP$$

since $G \in \mathcal{G}_1 \subset \mathcal{G}_2$.

[Note that $E[E[X \mid \mathcal{G}_1] \mid \mathcal{G}_2] = E[X \mid \mathcal{G}_1]$ a.s. by 6). If $\mathcal{G}_1 = \{\phi, \Omega\}$ and $\mathcal{G}_2 = \mathcal{G}$ then we get in particular from 7), $E[E[X \mid \mathcal{G}]] = EX$.]

8. (Jensen's inequality) Suppose $\varphi : \mathbb{R} \to \mathbb{R}$ is convex and $\varphi(X)$ is integrable, then

$$\varphi(E[X \mid \mathcal{G}]) \le E[\varphi(X) \mid \mathcal{G}] \text{ a.s.}$$

Proof: $\forall\ x_0 \in \mathbb{R}$, $\exists\ A(x_0)$ (= right derivative of φ at x_0) such that

$$\varphi(x_0) + A(x_0)(x - x_0) \le \varphi(x)\ \forall\ x \in \mathbb{R}.$$

Then

$$\varphi(E[X \mid \mathcal{G}]) + A(E[X \mid \mathcal{G}])(X - E[X \mid \mathcal{G}]) \le \varphi(X).$$

If X is bounded, all the terms are bounded - note that $x_0 \to A(x_0)$ is non-decreasing. Taking conditional expectations in the above inequality w.r.t. $\mathcal{G}$ we can verify 8). In the general case we can replace X by the bounded random variable $X_n = I_{G_n} X$ where $G_n = \{\omega : |X| \le n\}$. The proof is completed using a limiting argument.

9. Let $P : L^2_{\mathbb{R}}(\mathcal{F}) \to L^2_{\mathbb{R}}(\mathcal{G})$ be the orthogonal projection. Then $PX = E[X \mid \mathcal{G}]$ a.s.

Proof: PX satisfies, $EYX = EYP(X) \ \forall\ Y \in L^2(\mathcal{G})$.

Example 1 *suppose $X : \Omega \to \mathbb{R}^j$, $Y : \Omega \to \mathbb{R}^k$, X is $\mathcal{G}$ measurable and Y is independent of $\mathcal{G}$.*
a) Let $f(x) = P((x, Y) \in E)$ $E \in \mathcal{B}(\mathbb{R}^{k+\ell})$. Show that $P((X, Y) \in E \mid \mathcal{G}) = f(X)$ a.s. In particular

$$P(Y \in E' \mid \mathcal{G}) = P(Y \in E'), \ \text{a.s. } E' \in \mathcal{B}(\mathbb{R}^k).$$

b) $\varphi : \mathbb{R}^{k+j} \to \mathbb{R}$ is bounded measurable. Let $\varphi_1 : \mathbb{R}^j \to \mathbb{R}$ be defined as $\varphi_1(x) = E\varphi(x, Y)$. Show that $E(\varphi(X, Y) \mid \mathcal{G}) = \varphi_1(X)$. In particular $E(\psi(Y) \mid \mathcal{G}) = E\psi(Y)$ for any measurable $\psi : \mathbb{R}^k \to \mathbb{R}$ for which $\psi(Y)$ is integrable.

2 Discrete Parameter Martingales

$(\Omega, \mathcal{F}, P)$ a probability space.
$T = \{0, 1, 2, 3, \ldots\}$.
A family of sub σ-fields $\mathcal{F}_n, n \in T$ of $\mathcal{F}$ is called a *filtration* iff $m < n$ implies $\mathcal{F}_m \subset \mathcal{F}_n$. A collection of random variable $(X_n)_{n\in T}$ on $(\Omega, \mathcal{F}, P)$ is called a *stochastic process*. The process $(X_n)_{n\in T}$ is said to be *adapted* to the filtration $(\mathcal{F}_n)_{n\in T}$ iff $\forall\ n \in T$, X_n is $\mathcal{F}_n$ measurable.

Definition 2.1 *$\{X_n, \mathcal{F}_n\}_{n\in T}$ is said to be a* martingale *iff*

1. *$\forall\ n \in T, \ E|X_n| < \infty$.*
2. *(X_n) is $(\mathcal{F}_n)$ adapted.*
3. *$m < n \Rightarrow E(X_n \mid \mathcal{F}_m) = X_m$ a.s.*

$\{X_n, \mathcal{F}_n\}$ is said to be a sub-martingale (respectively super martingale) iff 1), 2) and 3′ hold where 3′ is given as follows:

3′. $m < n$ $E(X_n \mid \mathcal{F}_m) \geq X_m$ a.s. (respy. $\leq X_m$ a.s.).

Remark 2.2 *The process $(X_n)_{n\in T}$ is said to be a (sub, super) martingale iff $\{X_n, \mathcal{F}_n\}_{n\in T}$ is a (sub, super) martingale where $\mathcal{F}_n = \sigma\{X_u, u \leq n\}$.*

Example 2 *$X_0, X_1, X_2, \ldots, X_n$, independent $E|X_k| < \infty \ \forall \ k \geq 0, \quad EX_k = 0$. Define $S_n = X_0 + \ldots + X_n$ for $n \geq 0, \mathcal{F}_n = \sigma\{X_0, \ldots, X_n\}$*

$$\begin{aligned} E[S_{n+1}|\mathcal{F}_n] &= X_0 + \ldots + X_n + E[X_{n+1}|\mathcal{F}_n] \\ &= S_n. \end{aligned}$$

Thus $\{S_n, \mathcal{F}_n\}$ is a martingale. Note that $\{S_n, \mathcal{F}_n\}$ is a sub/super martingale if $EX_k \geq 0$ (respy. ≤ 0).

Example 3 *Suppose $p = (p_k)_{k\geq 0}$, is a sequence with $p_k \geq 0 \quad \sum_k p_k = 1$. Let $\{X_{n,i}, n \geq 0, i \in 1\}$ be i.i.d. random variables $P[X_{1,1} = k] = p_k$, $k \geq 0$. Let $Z_0 \equiv 1$ and define Z_k inductively as*

$$Z_{k+1} = \sum_{i=1}^{Z_k} X_{k,i} \qquad k = 0, 1, 2, \ldots$$

Let $\mathcal{F}_n = \sigma\{X_{k,i}, k < n, i \geq 1\}$. Then Z_n is $\mathcal{F}_n$ measurable.
Let $m = EX_{1,1}$ where we assume the expectation is finite. Define $W_n := m^{-n} Z_n, \quad n \geq 0$. Then $\{W_n, \mathcal{F}_n, n \geq 0\}$ is a martingale

$$\begin{aligned} E[W_{n+1}|\mathcal{F}_n] &= E\left[\sum_{i=1}^{Z_n} X_{n,i}|\mathcal{F}_n\right] \times m^{-(n+1)} \\ &= (m)^{-(n+1)} \left[E\sum_{i=1}^{k} X_{n,i}\right]_{k=Z_n} \\ &= (m)^{-(n+1)} \ Z_n \ EX_{1,1} \\ &= W_n. \end{aligned}$$

Example 4 *Let $D_1, D_2, \ldots, D_n, \ldots$ be $\{+1, -1\}$ valued random variables. Let $f_n : \mathbb{R}^{n-1} \times \{-1, +1\} \to \mathbb{R}$ be bounded and measurable for $n = 1, 2, 3, \ldots$. Let $x_0 \in \mathbb{R}$. Define the sequence of random variables $X_0, X_1, X_2, \ldots$ as follows: $X_0 \equiv x_0$.*

$$X_n := f_n(X_1, \ldots, X_{n-1}, D_n), \quad n = 1, 2, \ldots$$

Let $\mathcal{F}_n = \sigma\{X_0, X_1, \ldots, X_n\}, n = 0, 1, 2, \ldots$. An example of such a process is the C-R-R - model in finance (Cox-Ross-Rubinstein model). In such a process,

$$X_n = \begin{matrix} (1+b)X_{n-1} & \text{if } \ D_n = 1 \\ (1+a)X_{n-1} & \text{if } \ D_n = -1 \end{matrix}$$

where $a \in (-1, 0), b > 0$ *and* $\{D_n\}$ *is an i.i.d sequence* $P\{D_n = 1\} = 1 - P\{D_n = -1\} = p$, $p \in (0, 1)$*. If* $p = \frac{a}{a-b}$ *then it is easy to see that* $\{X_n, \mathcal{F}_n\}$ *is a martingale. Such martingales are called binary splitting martingales.*

Definition 2.3 *Let* $\tau : \Omega \to T \cup \{\infty\}$ *be a measurable map. We say that* τ *is a stopping time w.r.t. the filtration* $(\mathcal{F}_n)$ *iff* $(\tau \leq n) \in \mathcal{F}_n \ \forall \ n \in T$*. Note that* $\{\tau \leq n\} \in \mathcal{F}_\setminus \ \forall \ n \in T$ *iff* $\{\tau = n\} \in \mathcal{F}_n \ \forall \ n \in T$*.*

Facts

1. $\sigma \equiv n$ is a stop time.
2. If σ and τ are $(\mathcal{F}_n)$ stop times, so are $\sigma \wedge \tau$, $\sigma \vee \tau$.
3. For $c > 0$, $\sigma + c$ is an $(\mathcal{F}_n)$ stop time.
4. If $\sigma = \lim_k \sigma_k$ and $\sigma_k(\omega)$ increases to $\sigma(\omega)$ for every ω then σ is an $(\mathcal{F}_n)$ stop time.
5. If $\sigma = \lim_k \sigma_k$, $\sigma_k(\omega)$ decreases to $\sigma(\omega)$ for every ω then σ is an $(\mathcal{F}_n)$ stop time. $((\sigma < n) = \bigcup_k (\sigma_k < n))$.

Definition 2.4 *Let* σ *be an* $(\mathcal{F}_n)$ *stop time. The sigma field of events up to the random time* σ *viz.* $\mathcal{F}_\sigma$ *is defined as* $\mathcal{F}_\sigma = \{A \subset \Omega : \ \forall \ n \geq 0, A \cap \{\sigma \leq n\} \in \mathcal{F}_n\}$*.*

6. If $\sigma \leq \tau$ are two $(\mathcal{F}_n)$ stop times then $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

   Proof: Let $A \in \mathcal{F}_\sigma$.

   $$A \cap (\tau \leq n) = A \cap (\sigma \leq n) \cap (\tau \leq n).$$

7. Let $\sigma = \lim_k \sigma_k$ and suppose $\sigma_k(\omega)$ decreases to $\sigma(\omega)$ for every ω, then $\mathcal{F}_\sigma = \bigcap_k \mathcal{F}_{\sigma_k}$.

Proof: By 5) σ is an $\mathcal{F}_n$-stop time. By 6), $\mathcal{F}_\sigma \subset \bigcap_k \mathcal{F}_{\sigma_k}$. Let $A \in \bigcap_k \mathcal{F}_{\sigma_k}$. Then $A \cap (\sigma_k < n) \in \mathcal{F}_{n-1}$. Hence

$$A \cap (\sigma < n) = A \cap \left(\bigcup_k (\sigma_k < n) \right) \in \mathcal{F}_{n-1}.$$

It follows that $A \cap (\sigma \leq n) \in \mathcal{F}_n$ for every $n \geq 0$.

Definition 2.5 *A sequence of random variables* $(f_n)_{n\in T}$ *is said to be predictable w.r.t. the filtration* $(\mathcal{F}_n)_{n\in T}$ *iff* $\forall\ n \geq 1$, f_n *is* $\mathcal{F}_{n-1}$ *measurable and* f_0 *is a constant.*

Let $(X_n)_{n\in T}$ be a sequence of random variables and $(f_n)_{n\in T}$ be any other sequence of random variables. Define

$$\begin{aligned} Y_0 &\equiv X_0 \\ Y_n &= Y_{n-1} + f_n(X_n - X_{n-1}), \quad n = 1, 2, \ldots \\ &= Y_0 + \sum_{k=1}^{n} f_k(X_k - X_{k-1}). \end{aligned}$$

Proposition 2.6 *Suppose* $(f_n)_{n\in T}$ *is an* $(\mathcal{F}_n)$ *predictable sequence such that* $\exists\ c > 0$ *such that*

$$|f_n(\omega)| \leq c\ \forall\ n \in T,\ \forall\ \omega \in \Omega.$$

1. *If* $\{X_n, \mathcal{F}_n\}$ *is a martingale, then so is* $\{Y_n, \mathcal{F}_n\}$.
2. *Suppose* $f_n(\omega) \geq 0\ \forall\ n \geq 0, \omega \in \Omega$. *If* $\{X_n, \mathcal{F}_n\}$ *is a sub/super martingale then* $\{Y_n, \mathcal{F}_n\}$ *is also a sub/super martingale.*

Proof: 1) Since f_n is bounded, $f_n(X_n - X_{n-1})$ is integrable. Since (f_n) is a predictable sequence, f_n is $\mathcal{F}_{n-1}$ measurable.

$$\begin{aligned} E[f_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] &= f_n\ E(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) \\ &= 0 \\ \Rightarrow E[Y_n \mid \mathcal{F}_{n-1}] &= Y_{n-1} \text{ a.s.} \end{aligned}$$

2) If (X_n) is a submartingale and $0 \leq f_n(\omega) \leq c$ and (f_n) is $\mathcal{F}_n$-predictable then

$$\begin{aligned} E[f_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] &\geq 0 \text{ a.s.} \\ \Rightarrow \qquad E[Y_n \mid \mathcal{F}_{n-1}] &\geq Y_{n-1} \text{ a.s.} \end{aligned}$$

Theorem 2.7 *(Optional sampling theorem, Doob). Let* $(X_n, \mathcal{F}_n)$ *be a martingale. Let* σ *and* τ *be stopping time such that* $\sigma(\omega) \leq \tau(\omega) \leq c$ *for all* ω *for some integer* $c \geq 0$ *. Then* $E(X_\tau \mid \mathcal{F}_\sigma) = X_\sigma$ *a.s. In particular* $EX_\tau = EX_\sigma$.

Proof: Since $\{n \le \sigma\} = \left(\bigcup_{k=0}^{n-1} \{\sigma = k\}\right)^c \in \mathcal{F}_{n-1}$ the sequence $f_n = I_{\{n\le\sigma\}}$ is $(\mathcal{F}_n\}$ predictable. Let $(f \cdot X)_n \equiv Y_n$, the martingale in the previous proposition. Then

$$\begin{aligned}(f \cdot X)_n &= X_0 + \sum_{k=0}^{n} f_k(X_k - X_{k-1}) \\ &= \begin{cases} X_n & n \le \sigma \\ X_\sigma & n > \sigma \end{cases}\end{aligned}$$

i.e. $(f \cdot X_n) = X_{n\wedge\sigma}$.
In particular, if $\sigma \le c$ then

$$X_\sigma = X_{c\wedge\sigma} = (f \cdot X)_c \text{ is integrable.}$$

Similarly let $g_n = I_{\{n\le\tau\}}$. Then

$$\begin{aligned} X_\tau - X_\sigma &= (f \cdot X)_c - (g \cdot X)_c \\ \Rightarrow \quad EX_\tau - EX_\sigma &= E(f \cdot X)_c - E(g \cdot X)_c \\ &= EX_0 - EX_0 = 0. \end{aligned}$$

To show $E[X_\tau \mid \mathcal{F}_\sigma] = X_\sigma$ it suffices to show that

$$\int_A X_\sigma \, dP = \int_A X_\tau \, dP$$

for all $A \in \mathcal{F}_\sigma$. Let $A \in \mathcal{F}_\sigma$.
Define two stop times σ' and τ' as follows:

$$\begin{aligned} \sigma'(\omega) &= \sigma(\omega) \qquad \omega \in A \\ & \quad\ c \qquad \omega \in A^c \\ \tau'(\omega) &= \tau(\omega) \qquad \omega \in A \\ & \quad\ c \qquad \omega \in A^c. \end{aligned}$$

Then $\sigma'(\omega) \le \tau'(\omega) \le c$. Further if $k \ne c$ then $\{\sigma' = k\} = \{\sigma = k\} \cap A \in \mathcal{F}_k$. On the other hand if $k = c$, $\{\sigma' = c\}^c = \left(\bigcup_{k=1}^{c-1} \{\sigma = k\}\right) \cap A \in \mathcal{F}_c$. Thus σ' (and in a similar fashion) τ' are $(\mathcal{F}_k)$ stop times. Hence $EX_{\sigma'} = EX_{\tau'}$ and consequently

$$\begin{aligned} E1_A X_\sigma + E1_{A^c} X_c &= E1_A X_\tau + E1_{A^c} X_c \\ \Rightarrow \quad E1_A X_\sigma &= E1_A X_\tau. \end{aligned}$$

This completes the proof.

Corollary 2.8 *Let (X_n) be $(\mathcal{F}_n)$ adapted and integrable. Then $(X_n, \mathcal{F}_n)$ is a martingale iff for every pair of bounded stopping times σ, τ with $\sigma \le \tau$ we have*

$$EX_\tau = EX_\sigma.$$

Remark 2.9 *If (X_n) is an $(\mathcal{F}_n)$ submartingale and σ, τ as above then*

$$E[X_\tau \mid \mathcal{F}_\sigma] \ge X_\sigma \ \text{a.s.}$$

Similarly for super martingales with '$\le$' sign.

Theorem 2.10 *Let $(X_n, \mathcal{F}_n)$ be a submartingale. Let $\lambda > 0$ and $N \in T$.*
a)

$$\begin{aligned} \lambda P(\max_{0\le n\le N} X_n \ge \lambda) &\le E(X_N; \max_{0\le n\le N} X_n \ge \lambda) \\ &\le EX_N^+ \le E|X_N|. \end{aligned}$$

b)

$$\begin{aligned} \lambda P(\min_{0\le n\le N} X_n \le -\lambda) &\le -EX_0 + E(X_N : \min_{0\le n\le N} X_n > -\lambda) \\ &\le E|X_0| + E|X_N|. \end{aligned}$$

Proof: We first prove a). Define the stop time σ as follows:

$$\sigma = \begin{cases} \min\{n \le N : X_n \ge \lambda\} \\ N \text{ if the set in brackets viz. } \{\} \text{ is empty.} \end{cases}$$

It is easy to see that σ is a stop time: $\{\sigma = 0\} = \{X_0 \ge \lambda\}$ and if $1 \le k \le N-1$, then

$$\begin{aligned} \{\sigma = k\} &= \{X_1 < \lambda, \ldots X_{k-1} < \lambda, X_k \ge \lambda\} \cap \{X_0 < \lambda\} \\ \{\sigma = N\} &= \{X_1 < \lambda, \ldots X_{N-1} < \lambda, X_N \ge \lambda\} \cap \{X_0 < \lambda\} \\ &\quad \cup \{X_1 < \lambda, \ldots X_N < \lambda, X_0 < \lambda\} \end{aligned}$$

It follows that $\{\sigma = k\} \in \mathcal{F}_k \quad k = 0, 1, \ldots N$. Note that $\sigma \le N$. Hence by the Remark 2.9

$$\begin{aligned} EX_N &\ge EX_\sigma \\ &= E(X_\sigma : \max_{0\le n\le N} X_n \ge \lambda) \\ &\quad + E(X_N : \max_{0\le n\le N} X_n < \lambda) \\ &\ge \lambda P(\max_{0\le n\le N} X_n \ge \lambda) \\ &\quad + E(X_N : \max_{0\le n\le N} X_n < \lambda). \end{aligned}$$

This proves a).
b) Define

$$\tau = \begin{cases} \min\{n \leq N : X_n \leq -\lambda\} \\ N \text{ if the set in brackets viz. } \{\} \text{ is empty.} \end{cases}$$

$0 \leq \tau \leq N$ is a stop time as above. Hence $EX_0 \leq EX_\tau$ and the proof is completed in the same manner as above.

Corollary 2.11 *Let $(X_n, \mathcal{F}_n)$ be a martingale such that $E|X_n|^p < \infty$ for some $p \geq 1,\ n = 0, 1, 2, \ldots$.*
a) For every N,

$$P\left(\max_{0 \leq n \leq N} |X_n| \geq \lambda\right) \leq \frac{E|X_N|^p}{\lambda^p}.$$

b) If $p > 1$ then for every N,

$$E\left\{\max_{0 \leq n \leq N} |X_n|^p\right\} \leq \left(\frac{p}{p-1}\right)^p E|X_N|^p.$$

Proof: a) By Jensen's inequality $\{|X_n|^p, \mathcal{F}_n\}$ is a submartingale.

$$\begin{aligned} P\left\{\max_{0 \leq n \leq N} |X_n| \geq \lambda\right\} &= P\left\{\max_{0 \leq n \leq N} |X_n|^p \geq \lambda^p\right\} \\ &\leq \frac{E|X_N|^p}{\lambda^p}. \end{aligned}$$

b) Let $Y = \max_{0 \le n \le N} |X_n|$. From the (sub) martingale property,

$$\begin{aligned}
\lambda P(Y \ge \lambda) &\le \int I_{\{Y \ge \lambda\}} |X_N| \, dP \\
EY^p &= \int_\Omega dP \int_0^Y p \, \lambda^{p-1} d\lambda = \int_\Omega dP \int_0^\infty I_{\{Y \ge \lambda\}} p \, \lambda^{p-1} d\lambda \\
&= p \int_0^\infty d\lambda \, \lambda^{p-1} P(Y \ge \lambda) \\
&\le p \int_0^\infty d\lambda \, \lambda^{p-2} \int_\Omega I_{\{Y \ge \lambda\}} |X_N| \, dP \\
&= \frac{p}{p-1} \int_\Omega Y^{p-1} |X_N| \, dP \\
&\le \frac{p}{p-1} \, (E|X_N|^p)^{1/p} \, (E|Y|^p)^{1/p} \\
\Rightarrow \quad (E|Y|^p)^{1/p} &\le \frac{p}{p-1} (E|X_N|^p)^{1/p}.
\end{aligned}$$

Let (X_n) be any process and $a < b$. Define

$$\begin{aligned}
\tau_1 &= \min\{n : X_n \le a\} \\
\tau_2 &= \min\{n \ge \tau_1 : X_n \ge b\} \\
&\vdots \\
\tau_{2k+1} &= \min\{n \ge \tau_{2k} : X_n \le a\} \\
\tau_{2k+2} &= \min\{n \ge \tau_{2k+1} : X_n \ge b\} \\
&\vdots
\end{aligned}$$

Note that if (X_n) is $\mathcal{F}_n$-adapted thus each τ_i is an $(\mathcal{F}_n)$ stop time. This can be seen inductively as follows:

$$\{\tau_{2k+1} = i\} = \bigcup_{j=1}^{i-1} \{\tau_{2k} = j, X_{j+1} > a, \dots X_{i-1} > a, X_i \le a\}.$$

The right hand side $\in \mathcal{F}_i$ if $\{\tau_{2k} = j\} \in \mathcal{F}_j, j = 1 \dots i-1$. Define

$$U_N^X(a,b)(\omega) \equiv \max\{k : \tau_{2k} \le N\}.$$

We have the following important inequality:

Proposition 2.12 *(Upcrossing inequality). Let $(X_n, \mathcal{F}_n)$ be a submartingale. Then for every $N \geq 0$,*

$$EU_N(a,b) \leq \frac{1}{b-a}[E(X_N - a)^+ - E(X_0 - a)^+].$$

Proof: Since $Y_n = (X_n - a)^+$ is a submartingale, and $U_N^X(a,b) = U_N^Y(0, b-a)$, Without loss of generality we can take $a = 0$ and $b-a$ for b, and $X_n = Y_n$ in the definition of the stopping times τ_k. Let $\sigma_n \equiv \tau_n \wedge N$. If $2k > N$,

$$\begin{aligned} Y_N - Y_0 &= \sum_{n=1}^{2k} Y_{\sigma_n} - Y_{\sigma_{n-1}} \\ &= \sum_{n=1}^{k} Y_{\sigma_{2n}} - Y_{\sigma_{2n-1}} + \sum_{n=0}^{k-1} Y_{\sigma_{2n+1}} - Y_{\sigma_{2n}} \\ &\geq (b-a)U_N^Y(0, b-a) + \sum_{n=0}^{k-1} Y_{\sigma_{2n+1}} - Y_{\sigma_{2n}}. \end{aligned}$$

Note that $EY_{\sigma_{2n+1}} \geq EY_{\sigma_{2n}}$.
The Proposition is proved by taking expectations in the above inequality.

Theorem 2.13 *(Martingale Convergence). Let $(X_n, \mathcal{F}_n)$ be a submartingale such that $\sup_n EX_n^+ < \infty$. Then $X_\infty := \lim_{n\to\infty} X_n$ exists almost surely and $X_\infty \in L^1$.*

Proof: From the submartingale property of (X_n) we have

$$\begin{aligned} E|X_n| &= 2EX_n^+ - EX_n \\ &\leq 2EX_n^+ - EX_0. \end{aligned}$$

Note that if $M \leq N$ then $U_M(a,b)(\omega) \leq U_N(a,b)(\omega)$. Define $U(a,b)(\omega) = \lim_{N\to\infty} U_N(a,b)(\omega)$. Then using the hypothesis of the theorem, the Monotone convergence theorem and the inequality $E(X_N - a)^+ \leq EX_N^+ + |a|$

$$\begin{aligned} EU(a,b) &= \lim_{N\to\infty} EU_N(a,b) \\ &\leq \frac{1}{b-a} \limsup_{N\to\infty} [E(X_N - a)^+ - E(X_0 - a)^+] \\ &< \infty \end{aligned}$$

In particular, it follows that $U(a,b) < \infty$ a.s.

$$\{\omega : \underline{\lim}_{n\to\infty} X_n(\omega) < \overline{\lim}_{n\to\infty} X_n(\omega)\} = \bigcup_{\substack{a<b \\ a,b\in Q}} \{\omega : U(a,b) = \infty\}.$$

Consequently $P\{\underline{\lim}\ X_n < \overline{\lim}\ X_n\} = 0$ and the sequence X_n converges almost surely with a limit X_∞. Note that $X_\infty \in [-\infty, \infty]$. Using Fatou's lemma and the hypothesis of the theorem viz. $\sup\limits_n EX_n^+ < \infty$ it follows that

$$\begin{aligned} E|X_\infty| &\leq \underline{\lim}\ E|X_n| \\ &\leq \overline{\lim}\ (2EX_n^+ - EX_0) < \infty. \end{aligned}$$

3 Uniform Integrability

Definition 3.1 *A subset $\Lambda \subset L^1(\Omega, \mathcal{F}, P)$ is said to be uniformly integrable (U.I) iff*

$$\lim_{\lambda\to\infty} \sup_{X\in\Lambda} E(|X| : X > \lambda) = 0.$$

Proposition 3.2 *$\Lambda \subset L^1$ is uniformly integrable iff*
1) $\sup\limits_{X\in\Lambda} E|X| < \infty$.
2) $\lim\limits_{P(A)\to 0} \sup\limits_{X\in\Lambda} E(|X| : A) = 0$.

Proof: Let $\Lambda \subset L^1$ be uniformly integrable. Condition 1) follows easily from the definition of uniform integrability and the inequality

$$\int |X|\ dP \leq \int |X| + \lambda\{|X| > \lambda\}.$$

To prove the second condition, suppose that $\epsilon > 0$ is given. Choose $\lambda > 0$ such that

$$\sup_{X\in\Lambda} \int_{\{|X|>\lambda\}} |X|\ dP < \frac{\epsilon}{2}.$$

Let $0 < \delta < \frac{\epsilon}{2\lambda}$. Then if $P(A) < \delta$,

$$\begin{aligned}
\int_A |X|\, dP &= \int_{A\cap\{|X|>\lambda\}} |X|\, dP + \int_{A\cap\{|X|\le\lambda\}} |X| dP \\
&\le \int_{\{|X|>\lambda\}} |X|\, dP + \lambda P(A) \\
&< \frac{\epsilon}{2} + \lambda\delta < \epsilon.
\end{aligned}$$

The proof that the above two conditions imply uniform integrability is left as an exercise.

Proposition 3.3 *Suppose $\{X_n : n \ge 1\}$ is a sequence in L^1. Let $X_n \to X$ in probability. Then $X_n \to X$ in L^1 iff $\{X_n\}$ is uniformly integral.*

Proof: Suppose $X_n \in L^1$ and $X_n \to X$ in probability and $\{X_n\}$ is uniformly integrable. We note first that $X \in L^1$: This follows from Fatou's lemma and the fact that $X_n \to X$ in probability implies $X_{n_k} \to X$ a.s. for some subsequence $\{n_k\}$. For a random variable Y and $\lambda > 0$, let

$$Y^\lambda = Y \cdot I_{\{|Y|<\lambda\}}; \quad Y_\lambda = Y I_{\{|Y|\ge\lambda\}}.$$

Then

$$\|X_{n_k} - X\|_1 \le \|X^\lambda_{n_k} - X^\lambda\|_1 + \|X_{n_k\lambda}\|_1 + \|X_\lambda\|_1$$

where $\|Y\|_1 := E|Y|$. By the uniform integrability of $\{X_{n_k}\}$ and $\{X\}$ the last two terms go to zero (uniformly in n_k) for large λ. The first term goes to zero by the dominated convergence theorem, for fixed λ, as $k \to \infty$. This shows that any subsequence $\{n_k\} \subset \{n\}$ has a further subsequence $\{n'_k\} \subset \{n_k\}$ such that $X_{n'_k} \to X$ in L^1. It follows that $X_n \to X$ in L^1.
Suppose now that $X_n \to X$ in L^1. We show that $\{X_n\}$ is uniformly integrable: Given $\epsilon > 0$ choose $\delta > 0$ such that $\int_A |X|\, dP < \frac{\epsilon}{2}$ whenever $P(A) < \delta$. Also let $n_0 = n_0(\epsilon)$ be such that $\|X_n - X\|_1 < \frac{\epsilon}{2}$ for all $n \ge n_0$. Then, if $P(A) < \delta$ and $n \ge n_0$,

$$\begin{aligned}
\int_A |X_n|\, dP &\le \int_A |X|\, dP + \|X_n - X\|_1 \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Since $\{X_n\}_{n\geq n_0}$ is L^1 bounded, the uniform integrability of $\{X_n\}_{n\geq n_0}$ follows from the criterion for uniform integrability given in the previous proposition. Since a finite number of elements of L^1 is always uniformly integrable it follows that $\{X_n\}_{n\geq 1}$ is uniformly integrable.

Let $T' = T \cup \{\infty|] = \{0, 1, 2, \ldots\} \cup \{\infty\}$. We now consider the question of when a submartingale $\{X_n\} : n \in T$ can be extended as a submartingale to the set T'.

Theorem 3.4 *Let $\{X_n, \mathcal{F}_n, n \in T\}$ be a submartingale such that $\sup_n EX_n^+ < \infty$. Then $\{X_n, \mathcal{F}_n\}, n \in T'$ is a submartingale iff $\{X_n^+, n \geq 0\}$ is uniformly integrable.*

Proof: Suppose first that $\{X_n, \mathcal{F}_n\}, n \in T'$ is a submartingale. In particular we have $X_n \leq E(X_\infty \mid \mathcal{F}_n)$. By Jensen's inequality, a.s. $X_n^+ \leq E(X_\infty^+ \mid \mathcal{F}_n)$. Integrating over the set $\{X_n^+ > \lambda\}$ we get

$$E(X_n^+; X_n^+ > \lambda) \quad \leq \quad E(X_\infty^+; X_n^+ > \lambda).$$

On the other hand by Chebyshev's inequality and the assumption $\sup_n EX_n^+ < \infty$ we get

$$\lim_{\lambda\to\infty} \sup_n P(X_n^+ > \lambda) = 0.$$

Using the integrability of X_∞ we get

$$\lim_{\lambda\to\infty} \sup_n E(X_\infty^+; X_n^+ > \lambda) = 0.$$

It follows that

$$\lim_{\lambda\to\infty} \sup_n E(X_n^+; X_n^+ > \lambda) = 0.$$

Conversely, suppose $\{X_n^+, n \geq 0\}$ is U.I. From Proposition 3.3 it follows that $\lim_{n\to\infty} X_n = X_\infty$ in L^1 and a.s. Note that $X_n \to X_\infty$ a.s. Let $a > 0$. Note that $X_n \vee (-a) = (X_n + a)^+ - a$. By Jensen's inequality $\{X_n \vee (-a), \mathcal{F}_n\}$ is a submartingale . Since $|X_n \vee (-a)| \leq (X_n)^+ + |a|$, the submartingale $\{X_n \vee (-a), \mathcal{F})_n\}$ is uniformly integrable. Again by Proposition 3.3 it follows $X_n \vee (-a) \to X_\infty \vee (-a)$ in L^1 and a.s. By the dominated convergence theorem for conditional expectations

$$\begin{aligned} E(X_\infty \vee (-a) \mid \mathcal{F}_n) &= \lim_{m_k\to\infty} E(X_{m_k} \vee (-a) \mid \mathcal{F}_n) \text{ a.s.} \\ &\geq X_n \vee (-a) \text{ a.s.} \end{aligned}$$

Again using the dominated convergence theorem we can let $a \uparrow \infty$ in the above, to get

$$E(X_\infty \mid \mathcal{F}_n) \geq X_n \quad \text{a.s.}$$

In other words, $\{X_n, \mathcal{F}_n\}_{n \in T}$ is a submartingale.

Theorem 3.5 *Let $Y \in L^1$. Let $X_n = E[Y \mid \mathcal{F}_n]$ where $(\mathcal{F}_n)$ is any filtration. Then $\{X_n, \mathcal{F}_n\}$ is a uniformly integrable martingale, $X_\infty = \lim_{n\to\infty} X_n$ exists a.s. and in L^1. Define $\mathcal{F}_\infty := \sigma(\bigcup_n \mathcal{F}_n)$. Then*

$$X_\infty = E[Y \mid \mathcal{F}_\infty] \text{ a.s.}$$

Proof: Let $m < n$. Since $\mathcal{F}_m \subset \mathcal{F}_n$, by the property of conditional expectations we have

$$E[E[Y \mid \mathcal{F}_n] \mid \mathcal{F}_m] = E[Y \mid \mathcal{F}_m].$$

Thus $\{X_n, \mathcal{F}_n\}$ is a martingale. Since $|X_n| \leq E(|Y| \mid \mathcal{F}_n)$, the uniform integrability of $\{X_n\}$ follows as in the previous lemma. Since $\sup_n EX_n^+ < \infty$, $X_n \to X_\infty$ a.s. and hence in L^1. By the previous theorem, $\{X_n, \mathcal{F}_n, n \leq T'\}$ and $\{-X_n, \mathcal{F}_n, n \in T'\}$ are both submartingales. Hence $\{X_n, \mathcal{F}_n, n \in T'\}$ is a martingale. To show $X_\infty = E[Y \mid \mathcal{F}_\infty]$ a.s. suffices to show that for all $B \in \mathcal{F}_\infty$,

$$\int_B X_\infty = \int_B Y.$$

If $B \in \mathcal{F}_n \subset \bigcup_m \mathcal{F}_m$ the above equality follows since $\{X_n, \mathcal{F}_n, n \in T'\}$ is a martingale. The general case follows by a monotone class argument.

4 Reversed (Sub) Martingales

Let $(\mathcal{F}_{-n})_{n=0}^\infty$ be a decreasing sequence of sub σ-fields of $\mathcal{F}$ i.e. $\mathcal{F}_{-(n+1)} \subset \mathcal{F}_{-n} \subset \mathcal{F}$.

Definition 4.1 *The stochastic process $(X_{-n}, \mathcal{F}_{-n})_{n=0}^\infty$ is said to be a reversed (sub) martingale iff*
1) X_{-n} is $\mathcal{F}_{-n}$ measurable
2) $E|X_{-n}| < \infty \ \forall \ n$
3) $E(X_{-n} \mid \mathcal{F}_{-m}) \ (\geq) = X_{-m} \quad 0 \leq n \leq m$.

Remark 4.2 *Let $N > 0$. Let $Y_0 = X_{-N}, Y_1 = X_{-(N-1)} \dots Y_N = X_0$. $\mathcal{G}_0 = \mathcal{F}_{-N}$. $\mathcal{G}_1 = \mathcal{F}_{-(N-1)} \dots \mathcal{G}_N = \mathcal{F}_0$. Then $(X_{-n}, \mathcal{F}_{-n})_{n=0}^{\infty}$ is a reversed submartingale iff $\forall\ N \geq 0\ \ \{Y_n, \mathcal{G}_n\}_{n=0}^{N}$ is a submartingale.*

Example 5 *We denote by $S(n)$ the group of permutations on n elements. $\pi \in S(n)$ implies that $\pi : \{1, \dots n\} \to \{1, \dots n\}$ is 1-1 and onto.*
For $F : \mathbb{R}^n \to \mathbb{R}$ and $\pi \in S(n)$, define $F^{\pi} : \mathbb{R}^n \to \mathbb{R}$ by $F^{\pi}(x_1 \dots x_n) := F(x_{\pi(1)}, \dots, x_{\pi(n)})$. We say that F is n symmetric iff $F^{\pi} = F\ \ \forall\ \ \pi \in S(n)$.
Let $(X_n)_{n \geq 1}$ be a stochastic process. We define the σ-field of events invariant under permutations $\pi \in S(n)$ as follows:

$$\mathcal{E}_n \ := \ \sigma\{F(X_1, \dots X_n) : F : \mathbb{R}^n \to \mathbb{R}, \quad \textit{measurable}, F^{\pi} \equiv F \textit{ for all } \pi \in S(n)\},$$

Note that $\mathcal{E}_n$ is the correction of measurable subsets $A \subset \Omega$, such that there exists a Borel set $B \in \mathbb{R}^n$ satisfying $A = \{\omega : (X_{\pi_1} \dots X_{\pi_n}) \in B\}\ \forall\ \pi \in S(n)$.

Definition 4.3 *We say that the sequence of random variables $\{X_n\}_{n \geq 1}$ is exchangeable iff for all $n \geq 1$, $\pi \in S(n)$,*

$$(X_1 \dots X_n) \stackrel{d}{=} (X_{\pi(1)} \dots X_{(\pi(n))}).$$

Let $(X_n)_{n \geq 1}$ be an exchangeable sequence. Let $\mathcal{F}_{-n} \equiv \mathcal{E}_n, n \geq 1$. Define

$$Y_{-n} := \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Then $\{Y_{-n}, \mathcal{F}_{-n}, n \geq 1\}$ is an backwards martingale. This can be shown as follows: Let $A \in \mathcal{E}_n$. Suppose $A = \{(X_1 \dots X_n) \in B\}$. Let $F : \mathbb{R}^n \to \mathbb{R}$, $F \equiv I_B$. Then

$$\begin{aligned} EY_{-n+1} 1_A &= EY_{-n+1} F(X_1 \dots X_n) \\ &= EY^{\pi}_{-n+1} F(X_{\pi(1)}, \dots X_{\pi(n)}) \\ &= EY^{\pi}_{-n+1} 1_A \end{aligned}$$

where we have used the exchangeability of (X_n) in the second equality and the fact that $F^\pi = F$ in the last equality. Thus,

$$\begin{aligned} E\left[Y_{-n+1}|\mathcal{F}_{-n}\right] &= E\left[Y^\pi_{-n+1}|\mathcal{F}_{-n}\right] \\ &= \frac{1}{n!}\sum_{S\in S(n)} E\left[Y^\pi_{-n+1}|\mathcal{F}_{-n}\right] \\ &= \frac{1}{n!}\sum_{S\in S(n)} Y^\pi_{-(n-1)} \\ &= \frac{1}{n!}\sum_{S\in S(n)} \frac{1}{n-1}\left(X_{\pi(1)}+\ldots+X_{\pi(n-1)}\right) \\ &= Y_{-n}. \end{aligned}$$

Theorem 4.4 *Let $\{X_{-n}, \mathcal{F}_{-n}, n \geq 0\}$ be a reversed submartingale such that $\inf_n EX_{-n} > -\infty$. Then the sequence $\{X_{-n}\}$ is uniformly integrable and $X_{-\infty} := \lim_{n\to\infty} X_{-n}$ exists almost surely and in L^1.*

Proof: Since the number of upcrossings of (a, b) by $\{X_0, X_{-1}, \ldots X_{-(n-1)}, X_{-n}\}$ viz. $U^X_{-n}(a, b)$ differs by at most one from the number of upcrossings of (a, b) by $\{Y_0, \ldots, Y_n\}$ viz. $U^Y_n(a, b)$ where $Y_0 = X_{-n}, Y_1 = X_{-(n-1)}, \ldots, Y_n = X_0$. Therefore

$$\begin{aligned} EU^X_n(a,b) &\leq EU^Y_n(a,b) + 1 \\ &\leq \frac{1}{b-a}\left[E(X_0-a)^+ - E(X_{-n}-a)^+\right] + 1 \\ &\leq \frac{1}{b-a} E(X_0-a)^+ + 1 \end{aligned}$$

As in the proof of the martingale convergence theorem 2.13 it follows that $\lim_{n\to\infty} U^X_n(a, b) =: U^X_\infty(a, b) < \infty$ a.s. In particular, $X_{-\infty} := \lim_{n\to\infty} X_{-n}$ exists a.s. We now show that the sequence (X_{-n}) is uniformly integreable using the criterion of Proposition 3.2. To prove condition 1) of Proposition 3.2 note that

$$\begin{aligned} E|X_{-n}| &= 2\ EX^+_{-n} - EX_{-n} \\ &\leq 2\ EX^+_0 - \inf_n\ E\ X_{-n} \end{aligned}$$

It follows that $\sup_n\ E|X_{-n}| < \infty$ and hence by Fatou's lemma $E|X_{-\infty}| \leq \underline{\lim}_{n\to\infty}\ E|X_{-n}| < \infty$. We now show that condition 2) of Proposition 3.2

holds. Since $EX_{-(n+1)} \le EX_{-n}$, we have $\lim_{n\to\infty} EX_{-n} = \inf_n EX_{-n} > -\infty$.
Hence given $\epsilon > 0$, $\exists\ k \ge 1$ such that $EX_{-k} - \inf_n EX_{-n} < \epsilon$.
Further from the integrability of X_{-k}, there exists a $\delta > 0$, such that $E(|X_{-k}| : A) < \epsilon$ whenever $P(A) < \delta$. Then if $n \ge k$

$$\begin{aligned} E(|X_{-n}| : |X_{-n}| > \lambda) &= E(X_{-n}; X_{-n} > \lambda) - E(X_n; X_{-n} \le -\lambda) \\ &= E(X_{-n}; X_{-n} > \lambda) + E(X_{-n}; X_{-n} > -\lambda) - EX_{-n} \\ &\le E(X_{-k}; X_{-n} > \lambda) + E(X_{-k}; X_{-n} > -\lambda) - EX_{-k} + \epsilon \\ &= E(|X_{-k}|; |X_{-n}| > \lambda) + \epsilon \end{aligned}$$

Note that the probability of the set $\{X_{-n} > \lambda\}$ occurring in the first term on the right hand side of the last equality above can be made small for large λ, uniformly in n:

$$\begin{aligned} P(|X_{-n}| > \lambda) &\le \frac{1}{\lambda}\ E|X_{-n}| = \frac{1}{\lambda}(2EX^+_{-n} - EX_{-n}) \\ &\le \frac{1}{\lambda}(2EX^+_0 - \inf_n\ EX_{-n}) \\ &\le \delta \end{aligned}$$

if λ is chosen sufficiently large. This shows that condition 2) of Proposition 3.2 holds and completes the proof of the theorem.

References

[1] P. Billingsley. Probability and Measure. Wiley, India, 3rd edition, 2008.

[2] Achim Klenke. Probability Theory: A Comprehensive course. Springer-Verlag, Berlin, 2008.

[3] Yuan Shih Chow and Henry Teicher. Probability theory: Independence, interchangeability, martingales. Springer Texts in Statistics. Speringer-Verlag, New York, 3rd edition, 1997.

[4] Nobuyuki Ikeda and Shinzo Watanabe. Stochastic differential equations and diffusion processes, volume 24 of North-Holland Mathematical Library. North-Holland Publishing Co., Amsterdam, 2nd edition, 1989.

[5] Olav Kallenberg. Foundations of modern probability. Probabilitry and Its Applications. Springer-Verlag, New York and Berlin, 2nd edition, 2002.

[6] Ioannis Karatzas and Stevan E. Shreve. Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics, Springer-Verlag Berlin and Heidelberg GmbH & Co., 2nd edition, 2000.

[7] Paul-A. Maeyer. Probability and potentials. Blasidell Publishing Co. Ginn and Co., Waltham, Mass-Toronto, Ont-London, 1966.

[8] Albert N. Shiryaev. Probability, volume 95 of Graduate Texts in Mathematics. Springer-Verlag, New York, 2nd edition, 1996. Translation of the Russian edition from 1980.

[9] David Williams. Probability with martingales. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[10] K.R. Parthasarathy. Introduction to Probability and Measure. Texts and Readings in Mathematics, 33, Hindustan Book Agency, 2005.

[11] K.B. Athreya and S.N. Lahiri. Probability Theory, Trim Series, Vol.41, Hindustan Book Agency, New Delhi, 2007.