Solutions to the Selected Exercises in R. Durrett's Probability: Theory and Examples, 5th edition.

The 2rd part of those solutions can be find in this pdf file.

4.1.2. Prove Chebyshev's inequality. If $a > 0$ then

$$P(|X| \geq a|\mathscr{F}) \leq a^{-2}E[X^2|\mathcal{F}]$$

Proof: Notice that

$$X^2 \geq \mathbf{1}_{|X| \geq a}a^2, \quad \text{a.s..}$$

Therefore, form Theorem 4.1.9. (b) we have

$$E[X^2|\mathscr{F}] \geq a^2 E[\mathbf{1}_{|X|\geq a}|\mathscr{F}]. \quad \square$$

4.1.4. Use regular conditional probability to get the conditional Holder inequality from the unconditional one, i.e., show that if $p, q \in (1, \infty)$ with $1/p + 1/q = 1$ then

$$E[|XY||\mathcal{G}] \leq E(|X|^p|\mathcal{G})^{1/p}E(|Y|^q|\mathcal{G})^{1/q}.$$

Proof: Note that $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ is a nice space. Therefore, according to Theorem 4.1.17. there exists a μ which is the regular conditional distribution for (X, Y) given $\mathcal{G}$. In another word,

- For each $A \in \mathcal{B}(\mathbb{R}^2)$, $\omega \mapsto \mu(\omega, A)$ is a version of $P((X, Y) \in A|\mathcal{G})$.
- For a.e. w, $A \mapsto \mu(\omega, A)$ is a probability measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.

Now for a.e. w, it follows from the unconditional Holder inequality that

$$\int_{\mathbb{R}^d} |xy|\mu(\omega, dx, dy) \leq \Big(\int_{\mathbb{R}^2} |x|^p\mu(\omega, dx, dy)\Big)^{1/p}\Big(\int_{\mathbb{R}^2} |y|^q\mu(\omega, dx, dy)\Big)^{1/q}.$$

Using Theorem 4.1.16., the above inequality implies that

$$E[|XY||\mathcal{G}] \leq E(|X|^p|\mathcal{G})^{1/p}E(|Y|^q|\mathcal{G})^{1/q}, \quad \text{a.s.}$$

as desired. $\square$

4.1.6. Show that if $\mathcal{G} \subset \mathcal{F}$ and $EX^2 < \infty$ then

$$E(\{X - E(X|\mathcal{F})\}^2) + E(\{E(X|\mathcal{F}) - E(X|\mathcal{G})\}^2) = E(\{X - E(X|\mathcal{G})\}^2).$$

Proof: Note that

$$E(XE(X|\mathcal{F})) = E[E(XE(X|\mathcal{F})|\mathcal{F})] = E(E(X|\mathcal{F})^2).$$

Therefore we can verify the following Pythagorean law:

$$\begin{aligned} E(\{X - E(X|\mathcal{F})\}^2) &= E(X^2) + E(E(X|\mathcal{F})^2) - 2E(XE(X|\mathcal{F})) \\ &= E(X^2) - E(E(X|\mathcal{F})^2). \end{aligned}$$

Also note that $E[E(X|\mathcal{F})|\mathcal{G}] = E(X|\mathcal{G})$. So using this Pythagorean law three times, we get that the desired equality is equivalent to

$$E[X^2]-E[E(X|\mathcal{F})^2]+E[E(X|\mathcal{F})^2]-E[E(X|\mathcal{G})^2]=E[X^2]-E[E(X|\mathcal{G})^2].$$

This is trival. □

4.1.9. Show that if X and Y are random variables with $E(Y|\mathcal{G})=X$ and $EY^2=EX^2<\infty$, then $X=Y$ a.s.

Proof: Using the Pythagorean law (see the proof of exercise 4.1.6.), we have

$$E(\{Y-X\}^2)=E(Y^2)-E(X^2)=0$$

So we have $X=Y$ a.s. □

4.1.10. If $E|Y|<\infty$ and $E(Y|\mathcal{G})$ has the same distribution as Y, then $E(Y|\mathcal{G})=Y$ a.s.

Proof: First we proof that for each random variable X satisfies the condition of this exercise, we have $\{E(X|\mathcal{G})\geq 0\}\overset{a.s.}{=}\{X\geq 0\}$. In fact, on one hand, Jensen's inequality implies that

$$|E(X|\mathcal{G})|\leq E(|X||\mathcal{G}),\quad a.s..$$

On the other hand, the condition that $E(X|\mathcal{G})\overset{d}{=}X$ says that

$$E[|E(X|\mathcal{G})|]=E|X|=E(E(|X||\mathcal{G})).$$

So we must have $|E(X|\mathcal{G})|=E(|X||\mathcal{G})$ as surely.

This leads us to

$$\begin{aligned}E[X\mathbf{1}_{E[X|\mathcal{G}]\geq 0}]&=E[E[X|\mathcal{G}]\mathbf{1}_{E[X|\mathcal{G}]\geq 0}]=E[|E[X|\mathcal{G}]|\mathbf{1}_{E[X|\mathcal{G}]\geq 0}]\\&=E[E[|X||\mathcal{G}]\mathbf{1}_{E[X|\mathcal{G}]\geq 0}]=E[|X|\mathbf{1}_{E[X|\mathcal{G}\geq 0]}],\end{aligned}$$

which forces that

$$X\mathbf{1}_{E[X|\mathcal{G}]\geq 0}=|X|\mathbf{1}_{E[X|\mathcal{G}]\geq 0},\quad a.s..$$

So we must have $\{E(X|\mathcal{G})\geq 0\}\subset\{X=|X|\}=\{X\geq 0\}$. Noticing again we have

$$P(E(X|\mathcal{G})\geq 0)=P(X\geq 0),$$

so we must have $\{E(X|\mathcal{G})\geq 0\}\overset{a.s.}{=}\{X\geq 0\}$ as required. Now take $X=Y-c$, we get

$$\{E(Y|\mathcal{G})\geq c\}\overset{a.s.}{=}\{Y\geq c\},\quad c\in\mathbb{R}.$$

This complete the proof. □

4.2.2. Give an example of a submartingale X_n so that X_n^2 is a supermartingale.

Proof: $X_n=0$. □

4.2.3. Show that if X_n and Y_n are submartingales w.r.t. $\mathcal{F}_n$ then $X_n\vee Y_n$ is also.

Proof: Obviously $X_n\vee Y_n$ is adapted to $\mathcal{F}_n$. From $X_n\vee Y_n\leq|X_n|+|Y_n|$, we know $X_n\vee Y_n$ is integrable. Finally, we have

$$E[X_{n+1}\vee Y_{n+1}|\mathcal{F}_n]\geq E[X_{n+1}|\mathcal{F}_n]\vee E[Y_{n+1}|\mathcal{F}_n]\geq X_n\vee Y_n.\quad\square$$

4.2.4. Let $X_n, n\geq 0$, be a submartingale with $\sup X_n<\infty$. Let $\xi_n=X_n-X_{n-1}$ and suppose $E(\sup\xi_n^+)<\infty$. Show that X_n converges a.s..

Proof: For each $m \geq 0$, define stopping time $\tau_m := \inf\{k : X_k > m\}$. From $\sup_n X^+_{n\wedge\tau_m} \leq X^+_{\tau_m} = (X_{\tau_m-1} + \xi_{\tau_m})^+ \leq X^+_{\tau_m-1} + \xi^+_{\tau_m} \leq m + \sup_n \xi^+_n$, we know $E\sup_n X^+_{n\wedge\tau_m} < \infty$. According to Theorem 4.2.11, this says that $X_{n\wedge\tau_m}$ convergence a.s.. (Note that $X_{n\wedge\tau_m}$ is also a submartingale due to Theorem 4.2.9.). Therefore X_n convergence on the event $\{\tau_m = \infty\}$. Note that the condition $\sup X_n < \infty$ implies that $\bigcup_{m=1}^{\infty}\{\tau_m = \infty\} \overset{a.s.}{=} \Omega$. Therefore, X_n converges a.s.. $\square$

4.2.6. Let $(Y_k)_{k\in\mathbb{N}}$ be nonnegative i.i.d. random variables with $EY_m = 1$ and $P(Y_m = 1) < 1$. By example 4.2.3 that $X_n = \prod_{m\leq n} Y_m$ defines a martingale. (i) Show that $X_n \to 0$ a.s.. (ii) Use the strong law of large numbers to conclude $(1/n)\log X_n \to c < 0$.

Proof. (i) Since X_n is a non-negative martingale, it convergence a.s.ly to a limit, say X. Fix a $u > 0$ such that $P(|Y_\cdot - 1| \geq u) > 0$. Then for each $\epsilon > 0$, since X_n is independent of Y_{n+1}, we have

$$P(|X_{n+1} - X_n| \geq \epsilon u) \geq P(|X_n| \geq \epsilon)P(|Y_{n+1} - 1| \geq u).$$

The left hand side converges to 0 as $n \to \infty$, so we must have $P(|X_n| > \epsilon)$ converges to 0 as well. Therfore, $X = 0$ a.s..

(ii) We can assume that $P(Y_m = 0) = 0$, since if $P(Y_m = 0) > 0$, it is easy to see that

$$P(\exists n > 0 \text{ s.t. } X_n = 0) = 1,$$

which implies that $(1/n)\log X_n \to -\infty$.

Now, assuming $P(Y_m = 0) = 0$, we can write

$$\log X_n = \sum_{m=1}^{n} \log Y_1 \in (-\infty, \infty).$$

According to strong law of large numbers (Theorem 2.4.1. and 2.4.5.), we only have to show that $E\log Y_1 \in [-\infty, 0)$.

Define $Y_1^{(n)} = Y\mathbf{1}_{n^{-1}<Y<n}$, then both $Y_1^{(n)}$ and $\log Y_1^{(n)}$ are integrable. From Jensen's inequality, we have

$$E\log Y_1^{(n)} \leq \log EY_1^{(n)}, \quad n \geq 1.$$

By monotonicity, taking $n \to \infty$, we have

$$E\log Y_1 \leq \log EY_1 = 0.$$

Now, we only have to show that $E\log Y_1 \neq 0$. In fact, if $E\log Y_1 = 0$, we have $\log Y_1$ is integrable. So from $E\log Y_1 = \log EY_1$ and Exercise 1.6.1. we have $Y_1 = 1$ a.s., which contradicts to the condition $P(Y_1 = 1) < 1$. $\square$

4.2.8. Let X_n and Y_n be positive integrable and adapted to $\mathcal{F}_n$. Suppose

$$E(X_{n+1}|\mathcal{F}_n) \leq (1 + Y_n)X_n$$

with $\sum Y_n < \infty$ a.s.. Prove that X_n converges a.s. to a finite limit.

Proof: Let

$$W_n = \frac{X_n}{\prod_{m=1}^{n-1}(1 + Y_m)}, \quad n \in \mathbb{N},$$

which is positive, integrable and adapted to $\mathcal{F}_n$. From

$$E(W_{n+1}|\mathcal{F}_n) = \frac{1}{\prod_{m=1}^n (1+Y_m)} E(X_{n+1}|\mathcal{F}_n) \le W_n,$$

we know (W_n) is a supersomartingale. Theorem 4.1.12. says that (W_n) converges a.s. to a finite limit, say W. Since Y_m are positive, we have

$$\log \prod_{m=1}^n (1+Y_m) = \sum_{m=1}^n \log(1+Y_m) \le \sum_{m=1}^n Y_m.$$

From the condition $\sum Y_n < \infty$ a.s. and the fact that the left hand side of above is non-decreasing, we have $\prod_{m=1}^n (1+Y_m)$ converges a.s. to a finite limit. Therefore X_n also converges a.s. to a finite limit. ☐

4.2.9. Suppose X_n^1 and X_n^2 are supermartingales w.r.t. $\mathcal{F}_n$, and N is a stopping time so that $X_N^1 \ge X_N^2$. Then

$$Y_n = X_n^1 \mathbf{1}_{N>n} + X_n^2 \mathbf{1}_{N\le n}$$

and

$$Z_n = X_n^1 \mathbf{1}_{N\ge n} + X_n^2 \mathbf{1}_{N<n}$$

are supermartinales.

Proof: Clearely, Y_n and Z_n are integrable and adapted to $\mathcal{F}_n$. Note that

$$\begin{aligned} Y_{n+1} &= X_{n+1}^1 \mathbf{1}_{N>n+1} + X_{n+1}^2 \mathbf{1}_{N=n+1} + X_{n+1}^2 \mathbf{1}_{N<n+1} \\ &\le X_{n+1}^1 \mathbf{1}_{N>n+1} + X_{n+1}^1 \mathbf{1}_{N=n+1} + X_{n+1}^2 \mathbf{1}_{N<n+1} \\ &= Z_{n+1} = X_{n+1}^1 \mathbf{1}_{N>n} + X_{n+1}^2 \mathbf{1}_{N\le n}. \end{aligned}$$

Therefore,

$$E[Y_{n+1}|\mathcal{F}_n] \le E[Z_{n+1}|\mathcal{F}_n] = E[X_{n+1}^1|\mathcal{F}_n]\mathbf{1}_{N>n} + E[X_{n+1}^2|\mathcal{F}_n]\mathbf{1}_{N\le n} \le Y_n \le Z_n.$$

☐ **4.3.1.** Give an example of a martingale X_n with $\sup_n |X_n| < \infty$ and $P(X_n = a\ i.o.) = 1$ for $a = -1, 0, 1$.

Proof: Suppose that $(U_k)_{k\in\mathbb{N}}$ are i.i.d. r.v. with uniform distribution in $(0,1)$. Let $X_0 = 0$. For each $n \ge 1$, if $X_n = 0$, let $X_{n+1} = \mathbf{1}_{U_{n+1}\ge 1/2} - \mathbf{1}_{U_{n+1}<1/2}$; if $X_n \ne 0$, let $X_{n+1} = n^2 X_n \mathbf{1}_{U_{n+1}\le n^{-2}}$. Then (X_n) is a martingale since

$$E[X_{n+1}|\mathcal{F}_n] = \mathbf{1}_{X_n=0} E[X_{n+1}|\mathcal{F}_n] + \mathbf{1}_{X_n\ne 0} n^2 X_n E[U_{n+1} \le n^{-2}|\mathcal{F}_n] = X_n.$$

Note that

$$P(X_n > 1) = P(X_{n-1} \ne 0, U_n \le n^{-2}) \le \frac{1}{n^2}.$$

So B.C. lemma says that $P(X_n \le 1 \text{ for } n \text{ large engough}) = 1$, which says that $\sup_n |X_n| < \infty$ a.s..

It is elementary to see that

$$\sum_{n=1}^\infty P(X_{n+1} = 0|\mathcal{F}_n) = \sum_{n=1}^\infty P(X_{n+1} = 0|\mathcal{F}_n)\mathbf{1}_{X_n\ne 0} \overset{a.s.}{=} \sum_{n=1}^\infty (1 - \frac{1}{n^2})\mathbf{1}_{X_n\ne 0}.$$

Notice that we always have $\sum n^{-2}\mathbf{1}_{X_n\ne 0} < \infty$, so the above identity says that

$$\Big\{\sum_{n=1}^\infty P(X_{n+1} = 0|\mathcal{F}_n) < \infty\Big\} \overset{a.s.}{=} \Big\{\sum_{n=1}^\infty \mathbf{1}_{X_n\ne 0} < \infty\Big\} \subset \{X_n = 0\ i.o.\}.$$

Now, using Theorem 4.3.4. and above we have

$$\{X_n = 0\, i.o.\}^c \subset \{X_n = 0\, i.o.\}$$

in the sense of a.s.. This can only happen if $P(X_n = 0\, i.o.) = 1$.

We can also verify that

$$\sum_{n=1}^{\infty} P(X_{n+1} = 1|\mathcal{F}_n) \geq \sum_{n=1}^{\infty} P(X_{n+1} = 1|\mathcal{F}_n)\mathbf{1}_{X_n=0} = \frac{1}{2}\sum_{n=1}^{\infty}\mathbf{1}_{X_n=0}$$

So from what we have proved, we know that a.s.ly

$$\sum_{n=1}^{\infty} P(X_{n+1} = 1|\mathcal{F}_n) = \infty.$$

Using Theorem 4.3.4., we have that $P(X_n = 1\, i.o.) = 1$. Similarly, we have $P(X_n = -1,\ i.o.) = 1$. □

4.3.3. Let X_n and Y_n be positive integrable and adpted to $\mathcal{F}_n$. Suppose $E(X_{n+1}|\mathcal{F}_n) \leq X_n + Y_n$, with $\sum Y_n < \infty$ a.s.. Prove that X_n converges a.s. to a finite limit.

Proof: Define $M_n = X_n - \sum_{k=1}^{n-1} Y_k$. Then (M_n) is a supermartingale, since

$$E[M_{n+1}|\mathcal{F}_n] \leq X_n + Y_n - \sum_{k=1}^{n} Y_k = M_n.$$

Define stopping times

$$N_m := \inf\{n : \sum_{k=1}^{n} Y_k > m\}, \quad m \in \mathbb{N},$$

Then it is easy to see that, for each $m \in \mathbb{N}$,

$$M_{n\wedge N_m} + m = X_{n\wedge N_m} - \sum_{k=1}^{n\wedge N_m - 1} Y_k + m, \quad n \in \mathbb{N},$$

is a non-negative supermartingale. Therefore, M_n convergences a.s.ly on event $\{N_m = \infty\}$. Finally, notice that event

$$\bigcup_{m=1}^{\infty}\{N_m = \infty\} = \{\exists m \in \mathbb{N}\, s.t.\ \sum Y_n < m\}$$

is with probability 1. □

4.3.5. Show $\sum_{n=2}^{\infty} P(A_n| \cap_{m=1}^{n-1} A_m^c) = \infty$ implies $P(\cap_{m=1}^{\infty} A_m^c) = 0$.

Proof: Note that, there is a partition $\{\tilde{A}_n : n \in \mathbb{N}\}$ for the event $\bigcup_{m=1}^{\infty} A_m$ satisfying that

$$\bigcup_{m=1}^{n} A_m = \bigcup_{m=1}^{n} \tilde{A}_m, \quad n \in \mathbb{N}.$$

Define a filtration $(\mathcal{F}_n)$ such that

$$\mathcal{F}_n = \sigma(A_m; m = 1, \ldots, n) = \sigma(\tilde{A}_m; m = 1, \ldots, n), \quad n \in \mathbb{N}.$$

Notice also that $\{\tilde{A}_1, \ldots, \tilde{A}_n, \bigcap_{m=1}^{n} A_m^c\}$ is a partition for the underlying probablity space Ω. Therefore, according to Example 4.1.5., we have

$$P\Big(A_{n+1}\Big|\bigcap_{m=1}^{n} A_m^c\Big) = P(A_{n+1}|\mathcal{F}_n), \quad \text{on } \bigcap_{m=1}^{n} A_m^c.$$

From the condition of this exercise, we have

$$\sum_{m=1}^{\infty} P(A_{m+1}|\mathcal{F}_m) = \infty, \quad \text{on } \bigcap_{m=1}^{\infty} A_m^c.$$

Now, using Theorem 4.3.4. we get that

$$\bigcap_{m=1}^{\infty} A_m^c \subset \{A_m \; i.o.\} \subset \Big(\bigcap_{m=1}^{\infty} A_m^c\Big)^c$$

in the sense of almost sure. This can only happen if $P(\cap_{m=1}^{\infty} A_m^c) = 0$. □

For the next two exercises, in the context of Kakutani dichotomy for infinite product measures on page 235, suppose F_n, G_n are concentrated on $\{0,1\}$ and have $F_n(0) = 1-\alpha_n$, $G_n(0) = 1-\beta_n$.

4.3.9. Show that if $\sum \alpha_n < \infty$ and $\sum \beta_n = \infty$ then $\mu \perp \nu$.

Proof: Let $A := \{\xi_n \neq 0 \; i.o.\}$. According to B.C. lemma, condition $\sum \alpha_n < \infty$ says that $\mu(A) = 0$; condition $\sum \beta_n = \infty$ says that $\nu(A) = 1$. So we must have $\mu \perp \nu$. □

4.3.10. Suppose $0 < \alpha_n, \beta_n < 1$. Show that $\sum |\alpha_n - \beta_n| < \infty$ is sufficient for $\mu \ll \nu$ in general.

Proof: Let $(U_k)_{k\in\mathbb{N}}$ be i.i.d. r.v. uniform distribution on $[0,1]$ w.r.t. probability space $(\Omega, \mathcal{F}, P)$. Define $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ random element (ξ^1, ξ^2) by

$$\xi_k^1 := \mathbf{1}_{U_k \le \alpha_k}, \quad \xi_k^2 := \mathbf{1}_{U_k \le \beta_k}, \quad k \in \mathbb{N}.$$

Then we have ξ^1 has distribution μ and ξ^2 has distribution ν. Note that

$$\sum_{k=1}^{\infty} P(\xi_k^1 \neq \xi_k^2) = \sum_{k=1}^{\infty} |\alpha_k - \beta_k| < \infty,$$

therefore, according to B.C. lemma, we have

$$P\Big(\bigcup_{K=1}^{\infty} \bigcap_{k>K} \{\xi_k^1 = \xi_k^2\}\Big) = 1.$$

This says that there exists $K \in \mathbb{N}$ such that $P\big(\bigcap_{k>K}\{\xi_k^1 = \xi_k^2\}\big) > 0$. On the other hand, it is obvious that

$$P\Big(\bigcap_{k=1}^{K}\{\xi_k^1 = \xi_k^2\}\Big) > 0,$$

so from the independency, we have

$$P(\xi^1 = \xi^2) = P\Big(\bigcap_{k=1}^{K}\{\xi_k^1 = \xi_k^2\}\Big) \cdot P\Big(\bigcap_{k>K}\{\xi_k^1 = \xi_k^2\}\Big) > 0.$$

Now, suppose that $\mu \ll \nu$ is not true, then according to Kakutani dichotomy, we have $\mu \perp \nu$. This says that, there exists a subset $A \subset \mathbb{R}^{\mathbb{N}}$, wuch that $\mu(A) = \nu(A^c) = 1$. In this case, we have

$$P(\xi^1 \in A) = \mu(A) = 1 = \nu(A^c) = P(\xi^2 \in A^c),$$

which says that

$$P(\xi^1 \neq \xi^2) \geq P(\xi^1 \in A, \xi^2 \in A^c) = 1.$$

This is a contradiction. □

4.3.13. Galton and Watson who invented the process that bears their names were interested in the survival of family names. Suppose each family has exactly 3 children but coin flips determine their sex. In the 1800s, only male children kept the family name so following the male offspring leads to a branching process with $p_0 = 1/8$, $p_1 = 3/8$, $p_2 = 3/8$, $p_3 = 1/8$. Compute the probability ρ that the family name will die out when $Z_0 = 1$.

Proof: According to Theorem 4.3.12. we know that ρ is the only solution of

$$\varphi(\rho) = \rho$$

in $[0, 1)$, where

$$\varphi(\rho) = \frac{1}{8} + \frac{3}{8}\rho + \frac{3}{8}\rho^2 + \frac{1}{8}\rho^3.$$

Solving this gives that $\rho = \sqrt{5} - 2$. □ **4.4.3.** Suppose $M \leq N$ are stopping times. If $A \in \mathcal{F}_M$ then $L = M\mathbf{1}_A + N\mathbf{1}_{A^c}$ is a stopping time.

Proof: According to Theorem 7.3.6. we have $A^c \in \mathcal{F}_M \subset \mathcal{F}_N$. Therefore, for each $t \geq 0$, we have

$$A \cap \{M \leq t\} \in \mathcal{F}_t; \quad A^c \cap \{N \leq t\} \in \mathcal{F}_t.$$

From the above, we have $\{L \leq t\} = (A \cap \{M \leq t\}) \cup (A^c \cap \{N \leq t\}) \in \mathcal{F}_t$. □

4.4.5. Prove the following variant of the conditional variance formula. If $\mathcal{F} \subset \mathcal{G}$ then

$$E(E[Y|\mathcal{G}] - E[Y|\mathcal{F}])^2 = E(E[Y|\mathcal{G}])^2 - E(E[Y|\mathcal{F}]^2).$$

Proof: Note that $E\big(E[Y|\mathcal{G}]|\mathcal{F}\big) = E[Y|\mathcal{F}]$. So according to the Pythagorean law (see the Slution to Excise 4.1.6.) we get the desired result. □

4.4.7. Let X_n be a martingale with $X_0 = 0$ and $EX_n^2 < \infty$. Show that

$$P(\max_{1\leq m\leq n} X_m \geq \lambda) \leq \frac{EX_n^2}{EX_n^2 + \lambda^2}.$$

Proof: According to Theorem 4.2.6. we have $(X_n + c)^2$ is a submartingale where c is an arbitrary real number. Therefore, for each $c \in \mathbb{R}$, according to Doob's inequality

$$P(\max_{1\leq m\leq n} X_m \geq \lambda) \leq P\big(\max_{1\leq m\leq n} (X_n + c)^2 \geq (\lambda + c)^2\big) \leq \frac{E(X_n + c)^2}{(\lambda + c)^2} = \frac{EX_n^2 + c^2}{(\lambda + c)^2}.$$

Now, taking $c = \frac{EX_n^2}{\lambda}$ we have

$$P(\max_{1\leq m\leq n} X_m \geq \lambda) \leq \frac{EX_n^2}{EX_n^2 + \lambda^2}.$$

□

4.4.8. Let X_n be a submartingale and $\log^+ x = \max(\log x, 0)$. Prove

$$E\bar{X}_n \leq (1 - e^{-1})^{-1}\{1 + E(X_n^+ \log^+(X_n^+))\},$$

where $\bar{X}_n = \max_{k=1}^n X_k$.

Proof: Fix an $M > 1$. Note that

$$E[\bar{X}_n \wedge M] = \int_0^\infty P(\bar{X}_n \wedge M \geq \lambda)d\lambda \leq 1 + \int_1^M P(\bar{X}_n \wedge M \geq \lambda)d\lambda \leq 1 + \int_1^M P(\bar{X}_n \geq \lambda)d\lambda$$

Doob's inequality then says that

$$\begin{aligned} E[\bar{X}_n \wedge M] &\leq 1 + \int_1^M \frac{1}{\lambda} E[X_n^+; \bar{X}_n \geq \lambda]d\lambda \\ &\leq 1 + E[X_n^+ \int_1^M \frac{1}{\lambda} \mathbf{1}_{\bar{X}_n \geq \lambda} d\lambda] = 1 + E[X_n^+ \log(\bar{X}_n \wedge M)]. \end{aligned}$$

Now, use the calculus fact that $a \log b \leq a \log^+ a + b/e$, we have

$$E[\bar{X}_n \wedge M] \leq 1 + E[X_n^+ \log^+ X_n^+ + \frac{\bar{X}_n \wedge M}{e}].$$

This says that

$$E[\bar{X}_n \wedge M] \leq (1 - \frac{1}{e})^{-1}\{1 + E[X_n^+ \log X_n^+]\}.$$

Finally, taking $M \to \infty$, using monotone convergence theorem, we get the desired result. □

4.4.9. Let X_n and Y_n be martingales with $EX_n^2 < \infty$ and $EY_n^2 < \infty$. Show that

$$E[X_n Y_n] - E[X_0 Y_0] = \sum_{m=1}^n E[(X_m - X_{m-1})(Y_m - Y_{m-1})].$$

Proof: Since $E[X_{n+1} - X_n | \mathcal{F}_n] = 0$ and $Y_n \in \mathcal{F}_n$ we have by Theorem 4.4.7. that $E[(X_{n+1} - X_n)Y_n] = 0$. Similarly we have $E[(Y_{n+1} - Y_n)X_n] = 0$. Now it is easy to calculate that

$$\begin{aligned} &E(X_{n+1} - X_n)(Y_{n+1} - Y_n) = E(X_{n+1} - X_n)Y_{n+1} \\ &= E[X_{n+1}Y_{n+1} - X_n Y_n + X_n(Y_n - Y_{n+1})] = E[X_{n+1}Y_{n+1} - X_n Y_n]. \end{aligned}$$

From this to the desired result is trival. □

4.4.10. Let $X_n, n \geq 0$, be a martingale and let $\xi_n = X_n - X_{n-1}$ for $n \geq 1$. If $EX_0^2, \sum_{m=1}^\infty E\xi_m^2 < \infty$ then $X_n \to X_\infty$ a.s. and in L^2.

Proof: Using the result in Excise 4.4.9, we have

$$EX_n^2 = EX_0^2 + \sum_{m=1}^n E\xi_m^2.$$

Therefore $\sup_n EX_n^2 = EX_0^2 + \sum_{m=1}^\infty E\xi_m^2 < \infty$. According to Theorem 4.4.6. we get the desired result. □

4.6.4. Let X_n be r.v.'s taking values in $[0, \infty)$. Let $D = \{X_n = 0 \text{ for some } n \geq 1\}$ and assume

$$P(D | X_1, \ldots, X_n) \geq \delta(x) > 0 \quad a.s. \text{ on } \{X_n \leq x\}.$$

Use Theorem 4.6.9 to conclude that $P(D \cup \{\lim_n X_n = \infty\}) = 1$.

Proof: Let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n), n \geq 1$ and $\mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$. According to $D \in \mathcal{F}_\infty$, we have by Levy's 0-1 law that $E[D|\mathcal{F}_n] \to 1_D$ a.s.. For each $x > 0$, and each element $\omega \in \{X_n \leq x\, i.o.\}$, there exists a sequence of integers $(n_i)_{i \in \mathbb{N}}$ such that for each $i \in \mathbb{N}$, we have $X_{n_i}(\omega) \leq x$. Therefore, for this ω,

$$1_D(\omega) = \lim_n E[D|\mathcal{F}_n](\omega) = \lim_i E[D|\mathcal{F}_{n_i}](\omega) \geq \delta(x) > 0.$$

So we must have $1_D = 1$ on this event $\{X_n \le x \; i.o.\}$. This says that $\{X_n \le x \; i.o.\} \subset D$ for each $x > 0$. Therefore, $\cup_{x \in \mathbb{N}} \{X_m \le x \; i.o.\} \subset D$. Finally, noticing that $\cup_{x \in \mathbb{N}} \{X_m \le x \; i.o.\} = \{\lim_n X_n = \infty\}^c$, we must have the desired result. $\square$

4.6.5. Let Z_n be a branching process with offspring distribution p_k. Use the last result to show that if $p_0 > 0$ then $P(\lim_n Z_n = 0 \text{ or } \infty) = 1$.

Proof: Let $D := \{\lim_n Z_n = 0\} = \{Z_n = 0, \exists n \in \mathbb{N}\}$ be the event of extinction. Let $(\xi_i^n)_{i,n \in \mathbb{N}}$ be i.i.d. r.v. used in (4.3.4.). Let $\mathcal{F}_n = \sigma(Z_1, \ldots, Z_n)$. Now for each $x > 0$, on event $\{0 < Z_n \le x\}$ we have

$$P(D|\mathcal{F}_n) \ge P(Z_{n+1} = 0|\mathcal{F}_n) = P(\xi_i^{n+1} = 0, \forall i = 1 \ldots Z_n|\mathcal{F}_n) = p_0^{Z_n} \ge p_0^x > 0.$$

On event $\{Z_n = 0\}$, we have $P(D|\mathcal{F}_n) = 1 \ge p_0^x > 0$. Now, using Exercise 4.6.4. we have

$$P(D \cup \{\lim_n Z_n = \infty\}) = 1$$

as desired. $\square$

4.6.7. Show that if $\mathcal{F}_m \uparrow \mathcal{F}_\infty$ and $Y_n \to Y$ in L^1 then $E(Y_n|\mathcal{F}_n) \to E(Y|\mathcal{F}_\infty)$ in L^1.

Proof: According to Theorem 4.6.8. we have $E[Y|\mathcal{F}_n] \to E[Y|\mathcal{F}_\infty]$ in L^1. So we only have to show that

$$E(Y_n|\mathcal{F}_n) - E(Y|\mathcal{F}_n) \xrightarrow{L^1} 0.$$

In fact,

$$E\Big[\big|E(Y_n|\mathcal{F}_n) - E(Y|\mathcal{F}_n)\big|\Big] \le E\Big[E\big[|Y_n - Y|\big|\mathcal{F}_n\big]\Big] = E|Y_n - Y| \to 0. \quad \square$$

4.7.3. Prove directly from the definition that if $(X_l)_{l \in \mathbb{N}} \subset \{0, 1\}$ are exchangeable

$$P(X_1 = X_2 = \ldots X_k = 1|S_n = m) = \binom{n-k}{n-m} \Big/ \binom{n}{m}.$$

Proof: Define

$$\mathcal{N} = \Big\{w \in \{0,1\}^n : w_l = 1, \forall 1 \le l \le k; \sum_{l=1}^n w_l = m\Big\};$$
$$\mathcal{M} = \Big\{w \in \{0,1\}^n : \sum_{l=1}^n w_l = m\Big\}.$$

Note that, for each $w \in \mathcal{N}$, there exists a purmutation Γ_ω on $\{1, \ldots, n\}$ such that

$$w_{\Gamma_\omega(l)} = \mathbf{1}_{1 \le l \le m}, \quad l \in \{1, \ldots, n\}.$$

Now, writting $X = (X_l)_{l \in \mathbb{N}}$, we have

$$P(X_l = 1, 1 \le l \le k; S_n = m) = \sum_{w \in \mathcal{N}} P(X_l = w_l, \forall 1 \le l \le n) = \sum_{w \in \mathcal{N}} P(X_{\Gamma_\omega(l)} = \mathbf{1}_{1 \le l \le m}, \forall 1 \le l \le n)$$
$$= \sum_{w \in \mathcal{N}} P(X_l = \mathbf{1}_{1 \le l \le m}, \forall 1 \le l \le n) = \#\mathcal{N} \cdot P(X_l = \mathbf{1}_{1 \le l \le m}, \forall 1 \le l \le n).$$

Similarly we have $P(S_n = m) = \#\mathcal{M} \cdot P(X_l = \mathbf{1}_{1 \le l \le m}, \forall 1 \le l \le n)$.

Therefore, we have

$$P(X_1 = X_2 = \dots X_k = 1 | S_n = m) = \frac{\#\mathcal{N}}{\#\mathcal{M}} = \binom{n-k}{n-m} \Big/ \binom{n}{m}. \quad \square$$

4.7.4. If $(X_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ are exchangeable with $EX_i^2 < \infty$ then $E(X_1X_2) \geq 0$.

Proof: Note that

$$0 \leq \binom{n}{2}^{-1} E(X_1 + \dots + X_n)^2 = \binom{n}{2}^{-1} nEX_1^2 + EX_1X_2 \xrightarrow[n\to\infty]{} EX_1X_2. \quad \square$$

4.7.5. If $(X_k)_{k\in\mathbb{N}}$ are i.i.d. with $EX_i = \mu$ and $\mathrm{var}(X_i) = \sigma^2 < \infty$ then

$$\binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} (X_i - X_j)^2 \to 2\sigma^2.$$

Proof: Note that

$$\binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} (X_i - X_j)^2 =: A_n \in \mathcal{E}_n.$$

Therefore, since $\mathcal{E}_{n+1} \subset \mathcal{E}_n$, we know $\mathcal{F}_k := \mathcal{E}_{-k}, k = -1, -2, \dots$ is a filtration with index $-\mathbb{N}$. Therefore we have

$$\begin{aligned} A_n = E[A_n|\mathcal{E}_n] &= \frac{1}{(n)_2} \sum_{1\leq i,j\leq n} E[(X_i - X_j)^2|\mathcal{E}_n] = E((X_1 - X_2)^2|\mathcal{E}_n) \\ &= E((X_1 - X_2)^2|\mathcal{F}_{-n}) \xrightarrow[n\to\infty]{Thm\ 4.7.3.} E((X_1 - X_2)^2|\mathcal{F}_{-\infty}), \quad a.s. \end{aligned}$$

According to Hewitt-Savage 0-1 law, we have $\mathcal{F}_{-\infty} = \mathcal{E}$ is trival. So

$$E((X_1 - X_2)^2|\mathcal{F}_{-\infty}) = E((X_1 - X_2)^2) = 2\sigma^2. \quad \square$$

4.8.3. Let $S_n = \xi_1 + \dots + \xi_n$ where the ξ_i are independent with $E\xi_i = 0$ and $\mathrm{var}(\xi_i) = \sigma^2$. $S_n^2 - n\sigma^2$ is a martingale. Let $T = \min\{n : |S_n| > a\}$. Then we have $ET \geq a^2/\sigma^2$.

Proof: Without loss of generality, we assume $ET < \infty$. (Otherwise, the desired result is trival.) According to Wald's second identity (Excise 4.8.4. below), we have

$$\sigma^2 ET = ES_T^2 \geq a^2. \quad \square$$

4.8.4. Let $S_n = \xi_1 + \dots + \xi_n$ where the ξ_i are independent with $E\xi_i = 0$ and $\mathrm{var}(\xi_i) = \sigma^2$. Show that if T is a stopping time with $ET < \infty$ then $ES_T^2 = \sigma^2 ET$.

Proof: Since $S_{n\wedge T}^2 - (n \wedge T)\sigma^2, n \geq 1$ is a martingale, we have

$$E[S_{n\wedge T}^2 - \sigma^2(n \wedge T)] = 0, \quad n \in \mathbb{N}.$$

Therefore, we have

$$\sup_n E[S_{n\wedge T}^2] = \sigma^2 \sup_n E(n \wedge T) \leq \sigma^2 ET < \infty.$$

This tells us that $S_{n\wedge T}, n \geq 1$ is a L^2-martingale. Therefore $S_{n\wedge T} \xrightarrow{L^2} S_T$ and

$$ES_T^2 = \lim_{n\to\infty} ES_{n\wedge T}^2 = \lim_{n\to\infty} \sigma^2 E(n \wedge T) \overset{MCT}{=} \sigma^2 E(T). \quad \square$$

4.8.5. Let $(\xi_k)_{k\in\mathbb{N}}$ be independent with $P(\xi_i = 1) = p$ and $P(\xi_i = -1) = 1 - p$ where $p < 1/2$. Let $S_n = S_0 + \xi_1 + \dots + \xi_n$ and let $V_0 = \min\{n \geq 0 : S_n = 0\}$. Theorem 4.8.9 tells us that $E_xV_0 = x/(1 - 2p)$. Let $Y_i = \xi_i - (p - q)$ and note that $EY_i = 0$ and

$$\operatorname{var}(Y_i) = \operatorname{var}(X_i) = EX_i^2 - (EX_i)^2$$

then it follows that $(S_n - (p-q)n)^2 - n(1-(p-q)^2)$ is a martingale. (a) Use this to conclude that when $S_0 = x$ the variance of V_0 is

$$x \cdot \frac{1-(p-q)^2}{(q-p)^3}.$$

(b) Why must the answer in (a) be of the form cx?

Proof. (a). Since V_0 is a stopping time with finite expectation. Using Wald's second identity (Excise 4.8.7.), we have

$$E_x\left[(S_{V_0} - (p-q)V_0)^2 - V_0(1-(p-q)^2)\right] = x^2.$$

From the fact that $S_{V_0} = 0$ and $E_x V_0 = x/(1-2p)$, we can calculate the desired result.

(b) Define $V_y = \min\{n \geq 0, S_n = y\}$. Then, according to $S_0 = x > 0$, we have $V_x = 0$. From the fact that $|S_{n+1} - S_n| = 1$, and the fact that $EV_0 < \infty$, we know that

$$0 = V_x \leq V_{x-1} \leq \cdots \leq V_0 < \infty.$$

Moreover, it can be verified that $\{T_y = V_{y-1} - V_y : y = x, x-1, \ldots\}$ are i.i.d. random variables. (T_y are the time process (S_n) spend from first hitting position y to first hitting position $y-1$.) So

$$EV_0 = \sum_{k=1}^{x} ET_k = xc. \quad \square$$

4.8.7. Let S_n be a symmetric simple random walk starting at 0, and let $T = \inf\{n : S_n \notin (-a, a)\}$ where a is an integer. Find constants b and c so that $Y_n = S_n^4 - 6nS_n^2 + bn^2 + cn$ is a martingale and use this to compute ET^2.

Proof: First, since $S_n^2 - n$ is a martingale, we have $S_{n\wedge T}^2 - n \wedge T$ is a martinale. Therefore

$$E[S_{n\wedge T}^2] = E[n \wedge T].$$

Note that $S_{n\wedge T}^2$ is bounded by a^2; $n \wedge T$ is monotonic in n. Therefore, using bounded/monotonic convergence theorem, we get

$$a^2 = E[T]. \qquad (*)$$

It is elementary to verify that

$$E[Y_{n+1}|\mathcal{F}_n] - Y_n = (2b-6)n + b + c - 5.$$

Therefore, Y_n is a martingale iff $b = 3$ and $c = 2$. Now, set $b = 3, c = 2$. since $(Y_{T\wedge n})_{n\in\mathbb{N}}$ is also a martingale, we have

$$E[S_{n\wedge T}^4 + 3(n\wedge T)^2 + 2(n\wedge T)] = E[6(n\wedge T)S_{n\wedge T}^2].$$

Note that $(S_{n\wedge T}^4)_{n\in\mathbb{N}}$ is bounded by a^4; $(n\wedge T)_{n\in\mathbb{N}}$ is monotonic in n; $((n\wedge T)S_{n\wedge T}^2)_{n\in\mathbb{N}}$ is dominated by Ta^2. Therefore, using bounded/monotonic/dominated convergence theorem, we get

$$E[a^4 + 3T^2 + 2T] = E[6Ta^2].$$

From $(*)$, we have $ET^2 = (5a^4 - 2a^2)/3$. $\square$

4.8.10. Consider a favorable game in which the payoff ξ_k are $-1, 1$ or 2 with probability $1/3$ each. Use the results of the previous problem to compute the probability we ever go broke (i.e. our winings W_n reach 0) when we start with i.

Proof: It is elementary to verify that, if $\theta_0 = \ln(\sqrt{2}-1) < 0$, then

$$E[\exp(\theta_0 \xi_k)] = \frac{1}{3}(e^{-\theta_0} + e^{\theta_0} + e^{2\theta_0}) = 1.$$

It is well known that $X_n := \exp(\theta_0 W_n)$ is a martingale (the so-called exponential martingale). Note that it is non-negative, so it must have almost sure limit X_∞. In fact, since $W_n \to \infty$ almost surely, we must have $X_\infty = 0$.

Now, consider the martingale $X_{n\wedge T}$ where T is the broken time (hitting time at 0). Note that $W_{n\wedge T} \geq 0$, so $X_{n\wedge T} \in [0,1]$ is a bounded martingale. Therefore, we have

$$X_{n\wedge T} \xrightarrow[n\to\infty]{L^1} X_T \mathbf{1}_{T<\infty} + X_\infty \mathbf{1}_{T=\infty} = \mathbf{1}_{T<\infty}.$$

This implies that

$$P(T<\infty) = E[\mathbf{1}_{T<\infty}] = E[X_0] = \exp(\theta_0 i) = (\sqrt{2}-1)^i. \quad \square$$

6.1.1. Show that the class of invariant events $\mathcal{I}$ is a σ-field, and $X \in \mathcal{I}$ if and only if X is invariant, i.e., $X \circ \varphi = X$ a.s.

Proof: $\mathcal{I}$ is a sigma-field since (1) if $A \in \mathcal{I}$, then $\varphi^{-1}A^c = (\varphi^{-1}A)^c \overset{a.s.}{=} A^c$, which says that $A^c \in \mathcal{I}$. (2) $\emptyset \in \mathcal{I}$ since $\varphi^{-1}(\emptyset) = \emptyset$. (3) if $(A_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{I}$, then

$$\varphi^{-1}\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} \varphi^{-1}A_n \overset{a.s.}{=} \bigcup_{n\geq 1} A_n,$$

which says that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{I}$.

Also note that

$$\begin{aligned}
&X \in \mathcal{I} \\
\iff &\{X \in B\} \in \mathcal{I}, \forall \text{ Borel } B \\
\iff &\{\omega : X(\omega) \in B\} \overset{a.s.}{=} \varphi^{-1}\{\omega : X(\omega) \in B\} = \{\omega : X\circ\varphi(\omega) \in B\}, \forall \text{ Borel } B \\
\iff &X \overset{a.s.}{=} X \circ \varphi. \quad \square
\end{aligned}$$

6.1.2. Call A almost invariance if $P(A\Delta\varphi^{-1}(A)) = 0$ and call C invariant in the strict sense if $C = \varphi^{-1}(C)$. (i) Let A be any set, let $B = \cup_{n=0}^\infty \varphi^{-n}(A)$. Show $\varphi^{-1}(B) \subset B$. (ii) Let B be any set with $\varphi^{-1}(B) \subset B$ and let $C = \cap_{n=0}^\infty \varphi^{-n}(B)$. Show that $\varphi^{-1}(C) = C$. (iii) Show that A is almost invariant if and only if there is a C invariant in the strict sense with $P(A\Delta C) = 0$.

Proof: (i) $\varphi^{-1}B = \bigcup_{n=0}^\infty \varphi^{-1}\circ\varphi^{-n}A = \bigcup_{n=1}^\infty \varphi^{-n}A \subset B$. (ii) Since $\varphi^{-1}B \subset B$, we have that $\varphi^{-1}(C) = \bigcap_{n\in\mathbb{N}} \varphi^{-1}\varphi^{-n}B = \bigcap_{n=1}^\infty \varphi^{-n}B = \bigcap_{n=0}^\infty \varphi^{-n}B = C$.

(iii) Define B and C as above. Since A is invariance, we have $A \overset{a.s.}{=} \varphi^{-1}A$. It can be verified that if two measurable subsets Ω_1, Ω_2 of Ω satisfies $\Omega_1 \overset{a.s.}{=} \Omega_2$, then $\varphi^{-1}\Omega_1 \overset{a.s.}{=} \varphi^{-1}\Omega_2$. In fact,

$$P(\varphi^{-1}(\Omega_1)\Delta\varphi^{-1}(\Omega_2)) = P(\varphi^{-1}(\Omega_1\Delta\Omega_2)) = P(\Omega_1\Delta\Omega_2) = 0.$$

Using this fact multiple times we have $A \overset{a.s.}{=} \varphi^{-k}(A)$ for any $k \in \mathbb{N}$. Therefore $B \overset{a.s.}{=} A$. And we also have $B \overset{a.s.}{=} \varphi^{-k}(B)$ for any $k \in \mathbb{N}$. This tells us that $A \overset{a.s.}{=} C$. Yet (ii) already shows that C is strictly invariance. $\square$

6.1.3. (i) Show that if θ is irrational, $x_n = n\theta \bmod 1$ is dense in $[0,1)$. (ii) Use Theorem A.2.1. to show that if A is a Borel subset of $[0,1)$ with $|A| > 0$, then for any $\delta > 0$ there is an interval $J = [a,b)$ so that $|A \cap J| > (1-\delta)|J|$. (iii) Let θ be irrational. Combine this with (i) to conclude if A is an a subset of $[0,1)$ which is invariant under the operator

$$\varphi : y \mapsto y + \theta \bmod 1$$

and $|A| > 0$, then $|A| = 1$.

Proof: (i) Consider a 1-1 map $x \in [0,1) \mapsto e^{2\pi x i} \in S^1 := \{z \in \mathbb{C} : |z| = 1\}$. For any $\alpha, \beta \in S^1$, there is a natural distance

$$d(\alpha,\beta) := \text{length of the shorter arc connecting } \alpha \text{ and } \beta \text{ on } S^1.$$

We only need to prove that $\{\alpha_n = e^{2\pi n\theta i} : n \in \mathbb{N}\}$ is dense on S^1. More precisely, fixing an arbitrary β on S^1 and a large N, we only have to prove that there exists a n such that $d(\alpha_n, \beta) \leq \frac{2\pi}{N}$.

In fact, it is easy to verify that

1. $d(\alpha_n, \alpha_m) = d(0, \alpha_{n-m})$ for all $n, m \in \mathbb{N}$;
2. all α_n are distinct, so $d(\alpha_n, \alpha_m) \leq \frac{2\pi}{N}$ for some $m < n \leq N$. Fix this m and n.
3. $S^1 = \bigcup_{k=0}^{\infty} \{\alpha : \alpha \text{ lies on the shorter arc connecting } \alpha_{k(n-m)} \text{ and } \alpha_{(k+1)(n-m)}\}$

Now, for that fixed β, we know from 3. that there exists a $k \geq 0$ such that β lies on the shorter arc connecting $\alpha_{k(n-m)}$ and $\alpha_{(k+1)(n-m)}$. Therefore, for this k,

$$d(\alpha_{k(n-m)}, \beta) \leq d(\alpha_{k(n-m)}, \alpha_{(k+1)(n-m)}) = d(0, \alpha_{n-m}) \leq \frac{2\pi}{N},$$

as desired.

(ii) Let $\epsilon = \frac{\delta}{1-\delta}|A|$. Using Theorem A.2.1. there exists countable disjoint intervals $J_k := [a_k, b_k), k = 1, \ldots$ such that $A \subset \bigcup_{k=1}^{\infty} J_k$ and $\sum_{k=1}^{\infty} |J_k| < |A| + \epsilon$. Suppose that non of those intervals J_k satisfies the desired property that $|A \cap J| > (1-\delta)|J|$, then

$$\sum_{k=1}^{\infty} |J_k| \geq \sum_{k=1}^{\infty} \frac{|J_k \cap A|}{1-\delta} = \frac{|A|}{1-\delta}.$$

Therefore $\epsilon > \frac{|A|}{1-\delta} - |A| = \epsilon$. This is a contradiction.

(iii) Fix an arbitrary $1 > \delta > 0$. Note that if $J = [a,b)$ is an interval satisfies the condition $|A \cap J| \geq (1-\delta)|J|$ then either interval $J' = [a, \frac{a+b}{2})$ or interval $J'' = [\frac{a+b}{2}, b)$ also satisfy the same condition. This and (ii) implies that for any small $\epsilon > 0$, there exists an interval $J = [a,b)$ satisfies $|A \cap J| \geq (1-\delta)|J|$ and $|J| < \epsilon$. Fix this $\epsilon > 0$ and interval J. Let $N \in \mathbb{N}$ be the unique integer such that $\frac{1}{N+1} \leq |J| < \frac{1}{N}$. Thanks to (i), for each $k = 0, \ldots, N-1$, there exists an integer n_k such that

$$\frac{k}{N} \leq \varphi^{n_k} a < \varphi^{n_k} b < \frac{k+1}{N}.$$

Note that $J_k = [\varphi^{n_k} a, \varphi^{n_k} b), k = 0, \ldots N-1$ are disjoint intervals, A is φ-invariant i.e. $\varphi^{-1} A = A$, and φ is measure preserving. Therefore

$$
\begin{aligned}
|A| &\geq |\bigcup_{k=0}^{N-1}(A\cap J_k)| = \sum_{k=0}^{N-1}|A\cap J_k| = \sum_{k=0}^{N-1}|\varphi^{-n_k}(A\cap J_k)| \\
&= \sum_{k=0}^{N-1}|A\cap J| \geq N(1-\delta)|J| \geq (1-\epsilon)(1-\delta).
\end{aligned}
$$

Since $\epsilon > 0$ and $\delta > 0$ are arbitrary, so $|A| = 1$. $\square$

6.1.4. For any stationary sequence $\{X_n, n\geq 0\}$, there is a two-sided stationary sequence $\{Y_n : n\in\mathbb{Z}\}$ such that $(X_n)_{n\geq 0} \stackrel{d}{=} (Y_n)_{n\geq 0}$.

Proof: Give a stationary process $(X_n)_{n\in\mathbb{N}}$. According to Kolmogrove's extension theorem, there is a stochastic process $(Y_n)_{n\in\mathbb{Z}}$ such that for any $n_1 < n_2 < \cdots < n_k$, we have

$$
(Y_{n_i})_{i=1}^k \stackrel{d}{=} (X_{n_i - n_1})_{i=1}^k.
$$

(It is elementary to verify that hose finite dimensional distribution if consistent.) So, $(X_n)_{n\in\mathbb{N}} \stackrel{d}{=} (Y_n)_{n\in\mathbb{N}}$.

We also need to verify that $\{Y_n\}$ is stationary. This is elementary from its definition. $\square$

6.1.5. If $(X_k)_{k\in\mathbb{N}}$ is a stationary sequence and $g : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ is measurable then $Y_k = g(X_k, X_{k+1}, \ldots)$ is a stationary sequence. If X_n is ergodic then so is Y_n.

Proof: The shift operator is defined as usual

$$
\theta w = (w_2, w_3, \ldots), \quad w\in\mathbb{R}^{\mathbb{N}}.
$$

Define another operator $G : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ with

$$
Gw = (g(w), g(\theta w), \ldots, g(\theta^k w)), \quad w\in\mathbb{R}^{\mathbb{N}},
$$

then $Y = G(X)$. It can also be verified that $G\theta = \theta G$.

Therefore for each measurable subset $A\subset\mathbb{R}^{\mathbb{N}}$, we have

$$
\begin{aligned}
Y\in\theta^{-k}A \iff G(X)\in\theta^{-k}A \iff (\theta^k G)(X)\in A \iff (G\theta^k)(X)\in A \\
\iff \theta^k X\in G^{-1}A \iff X\in\theta^{-k}G^{-1}A.
\end{aligned}
$$

Therefore, if X is stationary, we have

$$
\mu_Y(\theta^{-1}A) = \mu_X(\theta^{-1}G^{-1}A) = \mu_X(G^{-1}A) = \mu_Y(A),
$$

which says that Y is also stationary.

Note that if A is invariant, i.e. $\theta^{-1}A = A$, then so is $G^{-1}A$, since $\theta^{-1}G^{-1}A = G^{-1}\theta^{-1}A = G^{-1}A$. Therefore, if X is egodic, then for any invariant subset $A\subset\mathbb{R}^{\mathbb{N}}$, we have

$$
\mu_Y(A) = \mu_X(G^{-1}A) \in \{0,1\},
$$

which says that Y is also ergodic. $\square$

6.1.6. Let $(X_k)_{k\in\mathbb{N}}$ be a stationary sequence. Let $n<\infty$ and let $(Y_k)_{k\in\mathbb{N}}$ be a sequence so that $(Y_{nk+1}, \ldots, Y_{n(k+1)}), k\geq 0$ are i.i.d. and $(Y_1, \ldots, Y_n) = (X_1, \ldots, X_n)$. Finally, let ν be uniformly distributed on $\{1, 2, \ldots, n\}$, independent of Y, and let $Z_m = Y_{\nu+m}$ for $m\geq 1$. Show that Z is stationary and ergodic.

Proof: The shift operator is defined as usual

$$
\theta w = (w_2, w_3, \ldots), \quad w\in\mathbb{R}^{\mathbb{N}}.
$$

It is easy to see that for each measurable $A \subset \mathbb{R}^{\mathbb{N}}$, we have

$$P(Y \in \theta^{-n} A) = P(Y \in A).$$

Therefore

$$P(Z \in \theta^{-1} A) = P(Y \in \theta^{-\nu-1} A) = \sum_{k=1}^{n} \frac{1}{n} P(Y \in \theta^{-k-1} A) = \sum_{k=1}^{n} \frac{1}{n} P(Y \in \theta^{-k} A) = P(Z \in A).$$

This says that Z is stationary.

Now assume that A is shift invariant i.e. $\theta^{-1} A = A$. Note that

$$\{Z \in A\} = \bigcup_{k=1}^{n} \{Z \in A, \nu = k\} = \bigcup_{k=1}^{n} \{Y \in \theta^{-k} A, \nu = k\} = \{Y \in A\}$$

Since Y is ergodic wrt operator θ^n and A is shift invariant wrt operator θ^n, so we have $P(Y \in A) \in \{0, 1\}$. This says that $P(Z \in A) \in \{0, 1\}$. So Z is ergodic. $\square$

6.1.7. Let $\varphi(x) = 1/x - [1/x]$ for $x \in (0, 1)$ and $A(x) = [1/x]$, where $[1/x] =$ the largest integer $\leq 1/x$. Then $a_n = A(\varphi^n x), n = 0, 1, 2, \ldots$ gives the continued fraction representation of x, i.e.

$$x = 1/(a_0 + 1/(a_1 + 1/\ldots)).$$

Show that φ preserves $\mu(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x}$ for $A \subset (0, 1)$.

Proof: It can be verified that for each $0 < a \leq b < 1$, we have

$$\varphi^{-1}[a, b) = \bigcup_{n \in \mathbb{N}} (\frac{1}{n+b}, \frac{1}{n+a}].$$

Therefore, we can calculate that

$$\mu\varphi^{-1}[a, b) = \sum_{n \in \mathbb{N}} \frac{1}{\log 2} \int_{\frac{1}{n+b}}^{\frac{1}{n+a}} \frac{dx}{1+x} = \frac{1}{\log 2} \int_a^b \frac{dx}{1+x} = \mu[a, b).$$

Using π-λ theorem, we can verifyopen φ preserves μ. $\square$

Exercise 6.2.1. Show tha if $X \in L^p$ with $p > 1$ then the convergence in Theorem 6.2.1 occurs in L^p.

Proof: Take an arbitrary $M > 0$. Let $X'_M := X\mathbf{1}_{|X| \leq M}$ and $X''_M := X\mathbf{1}_{|X| > M}$. We claim that

$$\limsup_{n \to \infty} \Big\| \frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k - E[X|\mathcal{I}] \Big\|_p \leq 2\|X''_M\|_p.$$

In fact, on one hand we have

$$\frac{1}{n} \sum_{k=0}^{n-1} X'_M \circ \varphi^k \xrightarrow[n \to \infty]{a.s. \& L^p} E[X'_M|\mathcal{I}],$$

where the almost sure convergence is due to Ergodic theorem, and the L^p convergence is then followed by bounded convergence theorem. On the other hand, we have

$$\begin{aligned} \Big\| \frac{1}{n} \sum_{k=1}^{n-1} X''_M \circ \varphi^k - E[X''_M|\mathcal{I}] \Big\|_p &\leq \frac{1}{n} \sum_{k=1}^{n-1} \|X''_M \circ \varphi^k\|_p + \|E[X''_M|\mathcal{I}]\|_p \\ &\leq 2\|X''_M\|_p, \quad n \geq 0. \end{aligned}$$

Now, since M is arbitrary and that $\|X''_M\|_p \to 0$ as $M \to \infty$, we get the desierd result. □

Exercise 6.2.2 (1) Show that if $g_n(w) \to g(w)$ a.s. and $E(\sup_k |g_k|) < \infty$, then $\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} g_m(\varphi^m w) = E(g|\mathcal{I}) \quad a.\,s.$

Proof: We claim that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} g_m \circ \varphi \leq E[g|\mathcal{I}] \quad \text{a.s..}$$

In fact, taking an arbitrary $M > 0$, we can define a almost sure finite random variable $h_M := \sup_{m\geq M} |g_m - g|$ using the condition $E(\sup_k |g_k|) < \infty$. Then we have by ergodic theorem that

$$\begin{aligned} \frac{1}{n} \sum_{m=0}^{n-1} g_m \circ \varphi^m &\leq \frac{1}{n} \sum_{m=0}^{M-1} (g + h_0) \circ \varphi^m + \frac{1}{n} \sum_{m=M}^{n-1} (g + h_M) \circ \varphi^m \\ &\xrightarrow[n\to\infty]{} E[g + h_M|\mathcal{I}] \quad a.\,s.\,. \end{aligned}$$

According to the fact that $|h_M| \leq g + \sup_k |g_k|$ and $h_M \to 0$ as $M \to \infty$, we have $E[g + h_M|\mathcal{I}] \to E[g|\mathcal{I}]$ a.s. due to the dominated convergence theorem. Therefore, the claim is true. Applying this claim to $(-g_n)_{n=1,...}$, we get that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} g_m \circ \varphi \geq E[g|\mathcal{I}] \quad \text{a.s..}$$

□

Exercise 6.2.3 Let $X_j = X \circ \varphi^j$, $S_k = X_0 + \cdots + X_{k-1}$, $A_k = S_k/k$ and $D_k = \max(A_1, \ldots, A_k)$. Show that if $\alpha > 0$ then

$$P(D_k > \alpha) \leq \alpha^{-1} E|X|.$$

Proof: Define $X'_j = X_j - \alpha$, $S'_k = X'_0 + \cdots + X_{k-1}$, $A'_k = S'_k/k$ and $D'_k = \max(A'_1, \ldots, A'_k)$. Then, it is easy to see that $S'_k = S_k - \alpha k$, $A'_k = A_k - \alpha$ and $D'_k = D_k - \alpha$. Lemma 6.2.2. says that $E[X'; D'_k > 0] \geq 0$. Therefore $E|X| \geq E[X; D_k > \alpha] \geq \alpha P(D_k > \alpha)$ as desired. □

Exercise 6.3.1 Let $g_n = P(S_1 \neq 0, \ldots, S_n \neq 0)$ for $n \geq 1$ and $g_0 = 1$. Show that

$$ER_n = \sum_{m=1}^{n} g_{m-1}$$

Where S_n and R_n is the same as Theorem 6.3.1.

Proof: Note that

$$R_n = 1 + 1_{S_{n-1} \notin \{S_n\}} + 1_{S_{n-2} \notin \{S_{n-1}, S_n\}} + \cdots + 1_{S_1 \notin \{S_2, \ldots, S_n\}}.$$

Therefore,

$$
\begin{aligned}
ER_n &= 1 + \sum_{m=1}^{n-1} P(S_m \notin \{S_{m+1}, \dots, S_n\}) \\
&= 1 + \sum_{m=1}^{n-1} P(S_{m+1} - S_m \neq 0, \dots, S_n - S_m \neq 0\}) \\
&= 1 + \sum_{m=1}^{n-1} P(S_1 \neq 0, \dots, S_{n-m} \neq 0\}) \\
&= \sum_{m=1}^{n} g_{m-1}.
\end{aligned}
$$

□

Exercise 6.3.2 Under the setting of Theorem 6.3.2. Show that if we assume $P(X_i > 1) = 0, EX_i > 0$, and the sequence X_i is ergodic, then $P(A) = EX_i$.

Proof: It is elementary analysis that if $s_n/n \to c > 0$, then we must have

$$n^{-1} \max_{1 \le k \le n} s_k \to c$$

and

$$\inf_{k=1,\dots} s_k > -\infty.$$

Ergodic theorem syas that

$$\frac{S_n}{n} \to EX_i > 0, \quad a.s.$$

so we must have

$$n^{-1} \max_{1 \le k \le n} S_k \to EX_i, \quad a.s.$$

and

$$M := \inf_{k=1,\dots} S_k > -\infty, \quad a.s.$$

Note, from the condition $P(X_i > 1) = 0$, we have

$$\max_{1 \le k \le n} S_k \le R_n \le \max_{1 \le k \le n} S_k - m,$$

which now implies that

$$\frac{R_n}{n} \to EX_i.$$

However, from Theorem 6.3.1. we already know that $n^{-1} R_n \to P(A)$. Therefore, we must have $P(A) = EX_i$. □

Exercise 6.3.3 Show that if $P(X_n \in A \text{ at least once}) = 1$ and $A \cap B = \emptyset$ then

$$E\Big(\sum_{1 \le m \le T_1} 1_{X_m \in B} \Big| X_0 \in A \Big) = \frac{P(X_0 \in B)}{P(X_0 \in A)}.$$

Proof: We can find a two-side stationary process which has the same finite demisional distribution same as $(X_n)_{n\in\mathbb{N}}$. With some abuse of notations, we denote such two-side stationary process as $(X_n)_{n\in\mathbb{Z}}$. Now, we can verify that

$$
\begin{aligned}
P(X_0 \in A) E\left[\left.\sum_{m=1}^{T_1} \mathbf{1}_{X_m \in B} \right| X_0 \in A\right] &= E\left[\sum_{t \in \mathbb{N}} \sum_{m=1}^{t} \mathbf{1}_{X_m \in B, T_1 = t}; X_0 \in A\right] \\
&= \sum_{m=1}^{\infty} P\left(X_m \in B, T_1 \geq m; X_0 \in A\right) = \sum_{m=1}^{\infty} P\left(X_0 \in A, X_1 \notin A, \ldots, X_{m-1} \notin A, X_m \in B\right) \\
&= \sum_{m=1}^{\infty} P\left(X_{-m} \in A, X_{-m+1} \notin A, \ldots, X_{-1} \notin A, X_0 \in B\right) = P(X_0 \in B).
\end{aligned}
$$

□

Exercise 6.3.4 Consider the special case in which $X_n \in \{0,1\}$, and let $\bar{P} = P(\cdot | X_0 = 1)$. Here $A = 1$ and so $T_1 = \inf\{m > 0 : X_m = 1\}$. Show $P(T_1 = n) = \bar{P}(T_1 \geq n)/\bar{E}T_1$.

Proof: From Theorem 6.3.3. we know that $P(X_0 = 1)\bar{E}T_1 = 1$. Therefore

$$
\frac{\bar{P}(T_1 \geq n)}{\bar{E}T_1} = P(T_1 \geq n | X_0 = 1) P(X_0 = 1) = P(T_1 \geq n, X_0 = 1).
$$

On the other hand, with some abuse of natations, assuming that $(X_n)_{n \in \mathbb{Z}}$ is a two-sided stationary sequence, we have

$$
\begin{aligned}
P(T_1 = n) &= \sum_{m=0}^{\infty} P(X_{-m} = 1, X_{-m+1} = 0, \ldots, X_{n-1} = 0, X_n = 1) \\
&= \sum_{m=0}^{\infty} P(X_0 = 1, X_1 = 0, \ldots, X_{m+n-1} = 0, X_{m+n} = 1) \\
&= P(X_0 = 1, T_1 \geq n).
\end{aligned}
$$

□