Section 9.1.

Exercise 9.1.13.

Solution.
Let (R,Θ) be the representation of (X,Y) in polar coordinates. Let $L:\mathbb{R}^+\times[0,2\pi)\to\mathbb{R}^2$ be the map defined by $L(r,\theta)=(r\cos(\theta),r\sin(\theta))$. For any $A\in\mathcal{B}_{\mathbb{R}^+\times[0,2\pi)}$,

$$\begin{aligned}\mathbb{P}\big((R,\Theta)\in A\big)&=\mathbb{P}\big(L(R,\Theta)\in L(A)\big)=\mathbb{P}\big((X,Y)\in L(A)\big)\\&=\int_{L(A)}\frac{1}{\pi}\mathbb{1}_{\{x^2+y^2\le 1\}}d\lambda(x,y)=\int_A\frac{1}{\pi}\mathbb{1}_{\{r^2\le 1\}}\left|\frac{\partial(x,y)}{\partial(r,\theta)}\right|d\lambda(r,\theta)\\&=\int_A\frac{r}{\pi}\mathbb{1}_{\{r\le 1\}}d\lambda(r,\theta).\end{aligned}$$

So the density of (R,Θ) is $\frac{r}{\pi}\mathbb{1}_{\{r\le 1\}}$. Now define $Z=\mathbb{1}_{\{Y\in A\}}$. Let $B\in\mathcal{B}_{[0,2\pi)}$, then by definition we have

$$\begin{aligned}\mathbb{E}\big(\mathbb{E}(Z|\Theta);\Theta\in B\big)&=\mathbb{P}(Y\in A,\Theta\in B)\\&=\mathbb{P}(R\sin(\Theta)\in A,\Theta\in B)\\&=\int_B\int_{[0,1)}\frac{r}{\pi}\mathbb{1}_A(r\sin(\theta))drd\theta\\&=\int_B p(1,\theta)d\theta,\end{aligned}$$

where $p(z,\theta)$ is the density of (Z,Θ). Then by **Exercise 9.1.10.**, the conditional expectation is computed as $g(\Theta)$, where

$$g(\theta)=\frac{p(1,\theta)}{p(1,\theta)+p(0,\theta)}=2\int_{[0,1)}r\mathbb{1}_A(r\sin(\theta))dr.$$

Hence $\mathbb{P}(Y\in A|\Theta)=2\int_0^1 r\mathbb{1}_A(r\sin(\Theta))dr$. □

Section 9.2.

Exercise 9.2.2.
If it's almost sure convergence in the statement, then it's not true in general.
Clarify. (Not converge almost surely) An example is sufficient. Let $\big([0,1],\mathcal{B}_{[0,1]},\mathbb{P}\big)$ be the probability space. Define random variables X_n as follows:

$$X_n=\mathbb{1}_{\left[\frac{k}{2^m},\frac{k+1}{2^m}\right]}\quad\text{if } n=2^m+k,\ k=0,\cdots,2^m-1.$$

Let $X=0$. Then $\|X_n-X\|_{L^1}=\|X_n\|_{L^1}=2^{-\lfloor\log_2(n)\rfloor}\to 0$ as $n\to\infty$, but $\{\omega\in[0,1]:\lim_{n\to\infty}X_n(\omega)=0\}=\emptyset$. Now take $\mathcal{G}=\mathcal{B}_{[0,1]}$. Of course X_n is $\mathcal{G}$-measurable and thus $\mathbb{E}(X_n|\mathcal{G})=X_n$ a.s. Hence we've shown that $X_n\to X$ in L^1 but $\mathbb{E}(X_n|\mathcal{G})$ doesn't converge to $\mathbb{E}(X|\mathcal{G})$ almost surely.

Solution. (Convergence in L^1) Suppose $X_n \to X$ in L^1. Then

$$\lim_{n\to\infty} \int_\Omega |X_n - X| d\mathbb{P} = \lim_{n\to\infty} \int_\Omega \mathbb{E}\left(|X_n - X||\mathcal{G}\right) d\mathbb{P} = 0.$$

By Jenson's inequality for conditional expectation, $\mathbb{E}\big(|X_n - X||\mathcal{G}\big) \geq \big|\mathbb{E}\left(X_n - X|\mathcal{G}\right)\big|$ a.e. for all $n \in \mathbb{N}$. Hence

$$\int_\Omega \mathbb{E}\left(|X_n - X||\mathcal{G}\right) d\mathbb{P} \geq \int_\Omega \big|\mathbb{E}\left(X_n - X|\mathcal{G}\right)\big| d\mathbb{P} \text{ for all } n \in \mathbb{N}.$$

Therefore,

$$0 = \lim_{n\to\infty} \int_\Omega \mathbb{E}\left(|X_n - X||\mathcal{G}\right) d\mathbb{P} \geq \limsup_{n\to\infty} \int_\Omega \big|\mathbb{E}\left(X_n - X|\mathcal{G}\right)\big| d\mathbb{P}.$$

Also, since $\liminf_{n\to\infty} \int_\Omega \big|\mathbb{E}\left(X_n - X|\mathcal{G}\right)\big| d\mathbb{P} \geq 0$, we have

$$\begin{aligned} 0 = \lim_{n\to\infty} \int_\Omega \big|\mathbb{E}\left(X_n - X|\mathcal{G}\right)\big| d\mathbb{P} &= \lim_{n\to\infty} \int_\Omega \big|\mathbb{E}\left(X_n|\mathcal{G}\right) - \mathbb{E}\left(X|\mathcal{G}\right)\big| d\mathbb{P} \\ &= \lim_{n\to\infty} \big\|\mathbb{E}(X_n|\mathcal{G}) - \mathbb{E}(X|\mathcal{G})\big\|_{L^1}, \end{aligned}$$

by the linearity of the conditional expectation. Hence $\mathbb{E}(X_n|\mathcal{G}) \to \mathbb{E}(X|\mathcal{G})$ in L^1. □

Exercise 9.2.3.

Solution. Let $\Omega = [0, 1]$. Define a collection of subsets $\mathcal{E}$ as

$$\mathcal{E} = \big\{\{Y \in O\} : O \text{ is open interval in } \Omega\big\}.$$

Claim. Let h be a measurable function such that $\mathbb{E}\big(\mathbb{E}(X|Y); A\big) = \mathbb{E}\big(h(Y); A\big)$ for all $A \in \mathcal{E}$, then $\mathbb{E}(X|Y) = h(Y)$ a.s.

Proof of claim. First note that $\mathcal{E}$ is a π-system since it is closed under finite intersections, and $\sigma(\mathcal{E}) = \sigma(Y)$. (Check: The collection $\mathcal{G} = \big\{B \subset \Omega : \{Y \in B\} \in \sigma(\mathcal{E})\big\}$ forms a sigma algebra. Clearly $\{O : O \text{ is open interval in } \Omega\} \subset \mathcal{G}$, hence $\mathcal{B}_{[0,1]} \subset \mathcal{G}$. Therefore $\sigma(Y) \subset \sigma(\mathcal{E})$.)

Suppose h is a measurable function that satisfies the condition. Since $\mathbb{E}(X|Y)$ and $h(Y)$ are non-negative measurable functions, $\mathbb{E}\big(\mathbb{E}(X|Y); \cdot\big)$ and $\mathbb{E}\big(h(Y); \cdot\big)$ induce two measures on Ω. Moreover, these two measures agree on a π-system $\mathcal{E}$. Now by a result of Dynkin's π-λ theorem (**Theorem 1.3.6.**), we can know that $\mathbb{E}\big(\mathbb{E}(X|Y); \cdot\big)$ and $\mathbb{E}\big(h(Y); \cdot\big)$ agree on $\sigma(\mathcal{E})$, hence agree on $\sigma(Y)$. Therefore $h(Y) = \mathbb{E}(X|Y)$ a.s. □

Given any open interval $(a, b) \in [0, 1]$. Fix $n \in \mathbb{N}$, then

$$\mathbb{E}\big(X; Y \in (a, b)\big) = \sum_{k=0}^{n-1} \int_{\frac{a+k}{n}}^{\frac{b+k}{n}} \omega d\omega = \sum_{k=0}^{n-1} \frac{(b+k)^2 - (a+k)^2}{2n^2}.$$

Now by the definition of Y, one can see that $\mathbb{P}(Y \leq y) = \sum_{k=0}^{n-1} \frac{y}{n} = y$ for $y \in [0, 1]$, hence

$Y \sim \text{Unif}([0,1])$. Now define $h : \Omega \to [0,1]$ as

$$h(y) = \sum_{k=0}^{n-1} \frac{y+k}{n^2}.$$

Then

$$\mathbb{E}\big(h(Y); Y \in (a,b)\big) = \int_a^b h(y) d\mu_Y(y) = \int_a^b \sum_{k=0}^{n-1} \frac{y+k}{n^2} dy = \sum_{k=0}^{n-1} \frac{(y+k)^2}{2n^2}\Big|_a^b,$$

which agrees on $\mathbb{E}(X; Y \in (a,b))$. Hence by the claim we have $\mathbb{E}(X|Y) = h(Y)$ a.s. $\square$

Exercise 9.2.4.

Solution. Since $\mathbb{E}(X_1; A) \le \mathbb{E}(X_2; A)$ for all $A \in \mathcal{F}$, $\mathbb{E}(X_1; X_1 > X_2) \le \mathbb{E}(X_2; X_1 > X_2)$. Then

$$\mathbb{E}(X_2; X_1 > X_2) - \mathbb{E}(X_1; X_1 > X_2) = \mathbb{E}(X_2 - X_1; X_1 - X_2 > 0) \ge 0.$$

This implies $\mathbb{P}(X_1 > X_2) = 0$. Hence $\mathbb{P}(X_1 \le X_2) = 1$, $\square$

Exercise 9.2.5. (TBD)

Solution. $\mathbb{E}\big[\big(X - \mathbb{E}(X|\mathcal{G})\big)\mathbb{E}(Y|\mathcal{G})\big] = 0$; $\mathbb{E}\big[\big(Y - \mathbb{E}(Y|\mathcal{G})\big)\mathbb{E}(X|\mathcal{G})\big] = 0$.

Exercise 9.2.6. (TBD)

Solution.

Exercise 9.2.7.

Solution. Fix $a \in \mathbb{N}$. Define $A = \{|X| \le a\}$. Then the random variables $X\mathbb{1}_A Y \mathbb{1}_{\{|Y| \le b\}}$ is bounded by $a|Y|$ for all $b \in \mathbb{N}$. Since $\mathbb{E}(a|Y|) = a\mathbb{E}|Y| < \infty$ and $X\mathbb{1}_A Y \mathbb{1}_{\{|Y| \le b\}} \to X\mathbb{1}_A Y$ as $b \to \infty$, by dominated convergence theorem we have

$$\begin{aligned}
\lim_{b\to\infty} \mathbb{E}(X\mathbb{1}_A Y \mathbb{1}_{\{|Y|\le b\}}) &= \mathbb{E}(X\mathbb{1}_A Y) \\
&= \mathbb{E}\big(\mathbb{E}(X\mathbb{1}_A Y|X)\big) \\
&= \mathbb{E}\big(X\mathbb{1}_A \mathbb{E}(Y|X)\big) \\
&= \mathbb{E}(X^2\mathbb{1}_A).
\end{aligned}$$

Since $\mathbb{E}(X\mathbb{1}_A Y) \le a\mathbb{E}|Y| < \infty$, $\mathbb{E}\big(X(X-Y); A\big) = 0$. Therefore,

$$\begin{aligned}
\mathbb{E}\big(X(X-Y); |X| < \infty\big) &= \mathbb{E}\Big(X(X-Y); \bigcup_{a\ge1}\{|X| \le a\}\Big) \\
&= \sum_{a\ge1} \mathbb{E}\big(X(X-Y); a < |X| \le a+1\big) \\
&\quad + \mathbb{E}\big(X(X-Y); |X| \le 1\big) \\
&= 0,
\end{aligned}$$

since $\mathbb{E}\big(X(X-Y);A\big) = 0$. Also note that $|X| < \infty$ a.s. since $\mathbb{E}|X| < \infty$, hence we have $\mathbb{E}\big(X(X-Y)\big) = 0$ a.s. Similarly one can prove $\mathbb{E}\big(Y(Y-X)\big) = 0$. Combine them together then we get $\mathbb{E}(X-Y)^2 = 0$, so $X - Y = 0$ a.s. $\square$

Exercise 9.2.8.

Solution. Let $U_t = \max\{S, t\}$. Then $\{U_t = t\} = \{S \le t\}$; $\{U_t > t\} = \{S > t\}$. Define

$$g(U_t) = \frac{\mathbb{E}(S; S \le t)}{\mathbb{P}(S \le t)} \mathbb{1}_{\{U_t = t\}} + U_t \mathbb{1}_{\{U_t > t\}}.$$

Then $\mathbb{E}\big(g(U_t); U_t = t\big) = \mathbb{E}(S; S \le t) = \mathbb{E}(S; U_t = t)$, and $\mathbb{E}\big(g(U_t); t < U_t < a\big) = \mathbb{E}(U_t; t < U_t < a) = \mathbb{E}(S; t < U_t < a)$. Hence $g(U_t) = \mathbb{E}(S|U_t)$ a.s. $\square$

Exercise 9.2.9.

Solution. (X, Y) can be wrote as a function of (Θ, Z):

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} \Theta & 1-\Theta \\ 1-\Theta & \Theta \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{2\Theta - 1} \begin{pmatrix} \Theta & \Theta - 1 \\ \Theta - 1 & \Theta \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$
$$= L(\Theta, Z).$$

Moreover, $Z \perp\!\!\!\perp \Theta$: Let μ be the law of (Z_1, Z_2). Let $C, D \in \mathcal{B}_{\mathbb{R}}$. Then

$$\mathbb{P}(Z \in C \times D) = \{\Theta = 1, X \in C, Y \in D\} \cup \{\Theta = 0, X \in D, Y \in C\},$$

hence by X, Y, Θ are independent and X, Y have identical distribution,

$$\begin{aligned}
\mathbb{P}(Z \in C \times D, \Theta = 1) &= p\mu(C)\mu(D) \\
&= p\Big(p\mu(C)\mu(D) + (1-p)\mu(D)\mu(C)\Big) \\
&= \mathbb{P}(\Theta = 1)\mathbb{P}(Z \in C \times D).
\end{aligned}$$

$\{\Theta = 0\}$ case is similar. Thus $Z \perp\!\!\!\perp \Theta$.

Now let $A \in \mathcal{B}_{\mathbb{R}^2}$. Then

$$\begin{aligned}
\mathbb{E}\big(g(X,Y); Z \in A\big) &= \int_\Omega g\big(L(\Theta, Z)\big) \mathbb{1}_{\{Z \in A\}} d\mathbb{P} \\
&= \int_{\mathbb{R}^2} \sum_{\theta = 0,1} g\big(L(\theta, \mathbf{z})\big)\nu(\theta) \mathbb{1}_A(\mathbf{z})\mu(\mathbf{z}) d\mathbf{z}, \quad (Z \perp\!\!\!\perp \Theta)
\end{aligned}$$

where ν is the law of Θ. Define h as follows,

$$h(\mathbf{z}) = \sum_{\theta=0,1} g\big(L(\theta, \mathbf{z})\big)\nu(\theta) = \mathbb{E}_\Theta\big((g \circ L)(\Theta, \mathbf{z})\big),$$

then $\mathbb{E}\big(g(X,Y); Z \in A\big) = \int_A h(\mathbf{z})\mu(\mathbf{z})d\mathbf{z}$. Hence

$$\mathbb{E}_\Theta\big((g \circ L)(\Theta, \mathbf{z})\big)\Big|_{\mathbf{z}=Z} = \mathbb{E}(g(X,Y)|Z) \text{ a.s.}$$

$\square$

Exercise 9.2.10.

Solution. Consider $B = \{\mathbb{E}(X_1|\mathcal{G}) \geq -b\}$. Then $\mathbb{E}(X_1|\mathcal{G})\mathbb{1}_B = \mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}) \geq -b$. Since $X_n \uparrow X$, for any $A \in \mathcal{G}$,

$$\mathbb{E}\big(\mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}); A\big) = \mathbb{E}(X_1\mathbb{1}_B; A) \leq \mathbb{E}(X_2\mathbb{1}_B; A) = \mathbb{E}\big(\mathbb{E}(X_2\mathbb{1}_B|\mathcal{G}); A\big).$$

But if $A = \{\mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}) > \mathbb{E}(X_2\mathbb{1}_B|\mathcal{G})\}$ then $\mathbb{E}\big(\mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}); A\big) \geq \mathbb{E}\big(\mathbb{E}(X_2\mathbb{1}_B|\mathcal{G}); A\big)$. Hence $\mathbb{E}\big(\mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}); A\big) = \mathbb{E}\big(\mathbb{E}(X_2\mathbb{1}_B|\mathcal{G}); A\big)$ a.s. This implies $A = \emptyset$, therefore $\mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}) \leq \mathbb{E}(X_2\mathbb{1}_B|\mathcal{G})$. Thus $\{\mathbb{E}(X_n\mathbb{1}_B|\mathcal{G})\}_{n\in\mathbb{N}}$ is monotonically increasing. Now use monotone convergence theorem of the nonnegative functions (since $b + \mathbb{E}(X_n\mathbb{1}_B|\mathcal{G}) \geq 0$), for any $A \in \mathcal{G}$,

$$\int_A \lim_{n\to\infty} \mathbb{E}(X_n\mathbb{1}_B|\mathcal{G})d\mathbb{P} = \lim_{n\to\infty} \int_A \mathbb{E}(X_n\mathbb{1}_B|\mathcal{G})d\mathbb{P} = \lim_{n\to\infty} \int_A X_n\mathbb{1}_B d\mathbb{P}.$$

Now we need to compute the last term. Since $0 \leq (X_n - X_1)\mathbb{1}_B \uparrow (X - X_1)\mathbb{1}_B$ a.s., by MCT again,

$$\lim_{n\to\infty} \int_A (X_n - X_1)\mathbb{1}_B d\mathbb{P} = \int_A (X - X_1)\mathbb{1}_B d\mathbb{P}.$$

Since $\mathbb{E}(X_1\mathbb{1}_B; A) = \mathbb{E}\big(\mathbb{E}(X_1\mathbb{1}_B|\mathcal{G}); A\big) \geq -b$, one can add this term in both sides. Then by the first equation,

$$\lim_{n\to\infty} \int_A X_n\mathbb{1}_B d\mathbb{P} = \int_A \lim_{n\to\infty} \mathbb{E}(X_n\mathbb{1}_B|\mathcal{G})d\mathbb{P} = \int_A X\mathbb{1}_B d\mathbb{P},$$

which is $\mathbb{E}\big(\mathbb{E}(X|\mathcal{G})\mathbb{1}_B; A\big)$. Since $A \in \mathcal{G}$ is arbitrary, by the first equation again, we get

$$\lim_{n\to\infty} \mathbb{E}(X_n|\mathcal{G})\mathbb{1}_B = \mathbb{E}(X|\mathcal{G})\mathbb{1}_B \text{ a.s.}$$

Since $B = \{\mathbb{E}(X_1|\mathcal{G}) \geq -b\}$ increases to $\cup_{b\geq 1}\{\mathbb{E}(X_1|\mathcal{G}) \geq -b\} = \{\mathbb{E}(X_1|\mathcal{G}) > -\infty\}$,

$$\lim_{n\to\infty} \mathbb{E}(X_n|\mathcal{G})\mathbb{1}_{\{\mathbb{E}(X_1|\mathcal{G})>-\infty\}} = \mathbb{E}(X|\mathcal{G})\mathbb{1}_{\{\mathbb{E}(X_1|\mathcal{G})>-\infty\}} \text{ a.s.}$$

$\square$

Exercise 9.2.11.

Solution. Fix $\epsilon > 0$. For each $n \in \mathbb{N}$, define X_n as

$$X_n = \sum_{k=0}^{n2^{n+1}-1} \Big(-n + \frac{k}{2^n}\Big)\mathbb{1}_{\{X\in[-n+\frac{k}{2^n}, -n+\frac{k+1}{2^n})\}}.$$

Then $\mathbb{E}|X - X_n|$ can be computed as:

$$
\begin{aligned}
\mathbb{E}|X - X_n| &= \mathbb{E}\Big|\sum_{k=0}^{n2^{n+1}-1}\Big(X + n - \frac{k}{2^n}\Big)\mathbb{1}_{\{X\in[-n+\frac{k}{2^n}, -n+\frac{k+1}{2^n})\}} + X\mathbb{1}_{\{X<-n,\, X\geq n\}}\Big| \\
&\leq \mathbb{E}\Big|\sum_{k=0}^{n2^{n+1}-1}\Big(X + n - \frac{k}{2^n}\Big)\mathbb{1}_{\{X\in[-n+\frac{k}{2^n}, -n+\frac{k+1}{2^n})\}}\Big| + \mathbb{E}\Big|X\mathbb{1}_{\{X<-n,\, X\geq n\}}\Big| \\
&\leq \sum_{k=0}^{n2^{n+1}-1}\mathbb{E}\Big|\Big(X + n - \frac{k}{2^n}\Big)\Big|\mathbb{1}_{\{X\in[-n+\frac{k}{2^n}, -n+\frac{k+1}{2^n})\}} + \mathbb{E}\big(|X|; X < -n,\, X \geq n\big) \\
&\leq \sum_{k=0}^{n2^{n+1}-1}\frac{1}{2^n}\mathbb{P}\Big(X \in \Big[-n + \frac{k}{2^n}, -n + \frac{k+1}{2^n}\Big)\Big) + \mathbb{E}\big(|X|; |X| \geq n\big) \\
&\leq \frac{1}{2^n} + \mathbb{E}\big(|X|; |X| \geq n\big).
\end{aligned}
$$

One can verify that the second term in the RHS goes to 0 as $n \to \infty$: Since $\mathbb{E}|X| < \infty$, $|X| < \infty$ a.e., therefore $|X|\mathbb{1}_{\{|X|\geq n\}} \downarrow 0$ as $n \to \infty$ a.e. Also, since $|X|\mathbb{1}_{\{|X|\geq n\}} \leq |X|$ for all n, then by dominated convergence theorem we have $\lim_{n\to\infty}\mathbb{E}(|X|; |X| \geq n) = 0$.

So there exists $N \in \mathbb{N}$ such that $\mathbb{E}|X - X_n| < \epsilon$ whenever $n > N$. Now suppose $n > N$. Since $\cup_{i=1}^{\infty}\mathcal{F}_i$ is the algebra that generates $\mathcal{F}$, there exists $A_{k+1} \in \cup_{i=1}^{\infty}\mathcal{F}_i$, $k = 0, \cdots, n2^{n+1} - 1$, such that

$$
\mathbb{P}\Big(A_{k+1}\Delta\Big\{X \in \Big[-n + \frac{k}{2^n},\, -n + \frac{k+1}{2^n}\Big)\Big\}\Big) < \frac{\epsilon}{n2^{k+1}}.
$$

(See Theorem 1.2.6.)

Then define the simple function Y_n as

$$
Y_n = \sum_{k=0}^{n2^{n+1}-1}\Big(-n + \frac{k}{2^n}\Big)\mathbb{1}_{A_{k+1}}.
$$

Then

$$
\begin{aligned}
\mathbb{E}|X_n - Y_n| &\leq \sum_{k=0}^{n2^{n+1}-1}\mathbb{E}\Big|\Big(-n + \frac{k}{2^n}\Big)\big(\mathbb{1}_{A_{k+1}} - \mathbb{1}_{\{X\in[\frac{k}{2^n}, \frac{k+1}{2^n})\}}\big)\Big| \\
&\leq \sum_{k=0}^{n2^{n+1}-1} n\mathbb{P}\Big(A_{k+1}\Delta\Big\{X \in \Big[-n + \frac{k}{2^n},\, -n + \frac{k+1}{2^n}\Big)\Big\}\Big) < \epsilon.
\end{aligned}
$$

Therefore $\mathbb{E}|X - Y_n| \leq \mathbb{E}|X - X_n| + \mathbb{E}|X_n - Y_n| < 2\epsilon$. $\square$

Fix any $\epsilon > 0$ again. There exists a $\mathcal{F}$-simple function Y_n, with sufficiently large n, such that $\mathbb{E}|X - Y_n| < \epsilon$. Since Y_n takes on finitely many sets in $\mathcal{F}$, there exists $N \in \mathbb{N}$ such that $A_1, \cdots, A_{n2^{n+1}} \in \mathcal{F}_m$ for any $m > N$. Clearly Y_n is $\mathcal{F}_m$-measurable. Therefore, if $m > N$,

$$
\begin{aligned}
\mathbb{E}\big|\mathbb{E}(X|\mathcal{F}_m) - X\big| &\leq \mathbb{E}\big|\mathbb{E}(X|\mathcal{F}_m) - Y_n\big| + \mathbb{E}\big|Y_n - X\big| \\
&\leq \mathbb{E}\big(\mathbb{E}(|X - Y_n||\mathcal{F}_m)\big) + \mathbb{E}|Y_n - X| \\
&= 2\mathbb{E}|X - Y_n| < 2\epsilon.
\end{aligned}
$$

Hence $\|\mathbb{E}(X|\mathcal{F}_m) - X\|_{L^1} \to 0$ as $m \to \infty$. $\square$

Exercise 9.2.12.

Solution. Since $\mathbb{E}(Y) < \infty$, $Y < \infty$ a.s. Hence $Y\mathbb{1}_{\{Y \geq a\}} \downarrow 0$ as $a \to \infty$ a.s. Also, $|Y\mathbb{1}_{\{Y \geq a\}}| \leq |Y|$ for all $a \in \mathbb{N}$. Then $\lim_{a\to\infty} \mathbb{E}(Y; Y \geq a) = 0$ by dominated convergence theorem.

Fix $\epsilon, \delta > 0$. Set $a \in \mathbb{N}$ to be the integer such that $\mathbb{E}(Y; Y \geq a) < \frac{\epsilon\delta}{3}$.

Since $0 \leq X_n \to 0$ in probability, we can take $N \in \mathbb{N}$ s.t. $\mathbb{P}(X_n > \frac{\epsilon\delta}{3}) < \frac{\epsilon\delta}{3a}$ whenever $n > N$. Suppose $n > N$. Since $X_n \geq 0$ a.s., $\mathbb{E}(X_n|\mathcal{G}) \geq 0$ a.s. Then by Markov's inequality,

$$\begin{aligned}
\mathbb{P}(\mathbb{E}(X_n|\mathcal{G}) > \epsilon) &< \frac{\mathbb{E}(\mathbb{E}(X_n|\mathcal{G}))}{\epsilon} = \frac{\mathbb{E}(X_n)}{\epsilon} \\
&= \frac{1}{\epsilon}\Big(\mathbb{E}\Big(X_n; X_n \leq \frac{\epsilon\delta}{3}\Big) + \mathbb{E}\Big(X_n; \frac{\epsilon\delta}{3} < X_n < a\Big) + \mathbb{E}(X_n; X_n \geq a)\Big) \\
&\leq \frac{1}{\epsilon}\Big(\frac{\epsilon\delta}{3} + a\mathbb{P}\Big(X_n > \frac{\epsilon\delta}{3}\Big) + \mathbb{E}(X_n; X_n \geq a)\Big) \\
&< \frac{1}{\epsilon}\Big(\frac{\epsilon\delta}{3} + \frac{\epsilon\delta}{3} + \mathbb{E}(X_n; X_n \geq a)\Big).
\end{aligned}$$

Since $0 \leq X_n \leq Y$ a.s. for all $n \in \mathbb{N}$, $\mathbb{E}(X_n; X_n \geq a) \leq \mathbb{E}(Y; X_n \geq a) \leq \mathbb{E}(Y; Y \geq a)$. Therefore

$$\mathbb{P}(\mathbb{E}(X_n|\mathcal{G}) > \epsilon) < \frac{1}{\epsilon}\Big(\frac{\epsilon\delta}{3} + \frac{\epsilon\delta}{3} + \mathbb{E}(Y; Y \geq a)\Big) < \delta.$$

Hence $\mathbb{P}(\mathbb{E}(X_n|\mathcal{G}) > \epsilon) \to 0$ as $n \to \infty$. $\square$

Exercise 9.2.13.

Solution. Let $\mathcal{P} = \{A \cap B : A \in \mathcal{G}_1, B \in \mathcal{G}_2\}$. Then $\mathcal{P}$ is a π-system: Pick $E_1, E_2 \in \mathcal{P}$, then $E_1 = A_1 \cap B_1$ and $E_2 = A_2 \cap B_2$ for some $A_1, A_2 \in \mathcal{G}_1$ and $B_1, B_2 \in \mathcal{G}_2$. Clearly $E_1 \cap E_2 \in \mathcal{P}$. Moreover, $\mathcal{G}_1 \cup \mathcal{G}_2 \subset \mathcal{P} \subset \sigma(\mathcal{G}_1 \cup \mathcal{G}_2) = \mathcal{G} \subset \sigma(\mathcal{P})$. Now let $A \cap B \in \mathcal{P}$. Then

$$\begin{aligned}
\mathbb{E}\big(\mathbb{E}(X|\mathcal{G}_1); A \cap B\big) &= \mathbb{E}\big(\mathbb{E}(X|\mathcal{G}_1)\mathbb{1}_A\mathbb{1}_B\big) \\
&= \mathbb{E}\big(\mathbb{E}(X|\mathcal{G}_1)\mathbb{1}_A\big)\mathbb{E}(\mathbb{1}_B) && (\mathcal{G}_1, \mathcal{G}_2 \text{ are independent}) \\
&= \mathbb{E}(X\mathbb{1}_A)\mathbb{P}(B) && (\text{conditional expectation on } \mathcal{G}_1) \\
&= \mathbb{E}(X\mathbb{1}_A\mathbb{1}_B) && (\sigma(\sigma(X) \cup \mathcal{G}_1), \mathcal{G}_2 \text{ are independent}) \\
&= \mathbb{E}\big(\mathbb{E}(X|\mathcal{G}); A \cap B\big). && (A \cap B \in \mathcal{G})
\end{aligned}$$

Hence $\mathbb{E}\big(\mathbb{E}(X|\mathcal{G}_1); \cdot\big) = \mathbb{E}\big(\mathbb{E}(X|\mathcal{G}); \cdot\big)$ agrees on $\mathcal{P}$. By a result of Dynkin's π-λ system, **Theorem 1.3.6.**, they agree on $\sigma(\mathcal{P})$ and hence on $\mathcal{G}$. Since $\mathbb{E}(X|\mathcal{G}_1)$ is of course a $\mathcal{G}$-measurable function, by definition $\mathbb{E}(X|\mathcal{G}_1) = \mathbb{E}(X|\mathcal{G})$ a.s. $\square$

Section 9.5.

Exercise 9.5.1.

Solution. $\{T_{a,b} = n\} = \bigcap_{i=1}^{n-1}\{a < S_i < b\} \cap \{S_n \in \{a, b\}\} \in \sigma(S_0, S_1, \cdots, S_n) = \mathcal{F}_n$.

□

Exercise 9.5.2.

Solution. Define $T = \min\{n : S_k - S_{k-1} = 1 \text{ for all } k = n - (a + b) + 1, \cdots, n\}$. Then $T_{a,b} < T$ a.s. And $\mathbb{E}T$ follows the following equation:

$$\mathbb{E}T = \sum_{k=0}^{a+b-1} p^k(1-p)(1 + k + \mathbb{E}T) + p^{a+b}(a + b)$$

Exercise 9.5.3.

Solution. ($T < \infty$ a.s.) One can check that $\max\{n : S_n \geq n\} \leq N$ if and only if $S_k < k$ for all $k > N$. Hence $\{T \leq N\} = \bigcap_{k>N}\{S_k < k\}$. Since X_n are i.i.d. and $\mathbb{E}X_n = 0$, by SLLN we know that $S_n/n \to 0$ almost surely. Therefore, $\Omega = \bigcup_{N\geq 1}\bigcap_{k>N}\{|S_k|/k < 1\}$ a.s. Since $\bigcap_{k>N}\{|S_k|/k < 1\} \subset \{T \leq N\}$, $\Omega = \bigcup_{N\geq 1}\{T \leq N\} = \{T < \infty\}$ a.s.

(T is not a stopping time) Note that $\{T = 0\} = \bigcap_{k\geq 1}\{S_k < k\}$. Clearly $\{T = 0\} \neq \emptyset$, since the path $\omega = (0, -1, -2, -3, \cdots) \in \bigcap_{k\geq 1}\{S_k < k\}$. Suppose $\{T = 0\} = \Omega$. Then $\{S_1 < 1\} = \{X_1 < 1\} = \Omega$. But $\mathbb{P}(X_1 < 1) = 1/2 \neq 1$, a contradiction. So $\{T = 0\} \notin \{\emptyset, \Omega\} = \mathcal{F}_0$, therefore T is not a stopping time.

□

Exercise 9.5.4.

Solution. Suppose T is a stopping time. Then $\{T \leq n\} = \{T > n\}^c \in \cup_{i=0}^{n}\mathcal{F}_i = \mathcal{F}_n$, so $\{T > n\} \in \mathcal{F}_n$. Suppose $\{T > n\} \in \mathcal{F}_n$ for all $n \geq 0$. Then $\{T \leq n + 1\} \in \mathcal{F}_{n+1}$ and $\{T \geq n+1\} = \{T > n\} \in \mathcal{F}_n \subset \mathcal{F}_{n+1}$, hence $\{T = n+1\} = \{T \leq n+1\} \cap \{T \geq n+1\} \in \mathcal{F}_{n+1}$. On the other hand, $\{T = 0\} = \{T \leq 0\} = \{T > 0\}^c \in \mathcal{F}_0$. So $\{T = n\} \in \mathcal{F}_n$ for all $n \geq 0$.

□

Exercise 9.5.5.

Solution. Let $k \in \mathbb{N}$. Then $\{T = k\} \cap \{T = n\} = \emptyset \in \mathcal{F}_n$ if $k \neq n$, and $\{T = k\} \cap \{T = n\} = \{T = n\} \in \mathcal{F}_n$ if $k = n$. Hence $\{T = k\} \in \mathcal{F}_T$. So T is $\mathcal{F}_T$-measurable.

□

Exercise 9.5.6.

Solution. Consider a filtration $\mathcal{F}_n = \sigma(S_0, S_1, \cdots, S_n)$ as in **Exercise 9.5.1.** Define $T = \min\{n : S_n = 2\}$. Let $A = \{S_1 = 1, S_2 = 0, S_3 = 1, S_4 = 2\}$. Clearly A is a proper subset of $\{T = 4\}$, and hence $A \notin \sigma(T)$. But $A \cap \{T = j\} = \emptyset \in \mathcal{F}_n$ for all $j \neq 4$, and $A \cap \{T = 4\} = A \in \mathcal{F}_4$. Thus

$A \in \mathcal{F}_T \setminus \sigma(T)$. By the previous exercise we then know that $\sigma(T)$ is a proper subset of $\mathcal{F}_T$.

□

Exercise 9.5.7.

Solution. Let $n \geq 0$. Then $A \cap \{T_{a,b} = n\} = \{\text{visits } 0 \text{ exactly } k \text{ times by time } n\} \cap \{T_{a,b} = n\} \in \mathcal{F}_n$, since the first set is in $\sigma(S_0, \cdots, S_n) = \mathcal{F}_n$. So by definition $A \in \mathcal{F}_T$.

□

Exercise 9.5.8.

Solution. Let $E \in \mathcal{B}_{\mathbb{R}}$. $X_T(\omega) \in E$ if and only if $T(\omega) = n$ for some n and $X_n(\omega) \in E$. Hence $\{X_T \in E\} = \cup_{n \geq 0} \{X_n \in E,\, T = n\}$. Therefore for any $k \geq 0$,

$$\{X_T \in E\} \cap \{T = k\} = \{X_k \in E,\, T = k\} \in \mathcal{F}_k.$$

So $\{X_T \in E\} \in \mathcal{F}_T$, i.e. X_T is $\mathcal{F}_T$-measurable.

□

9.6.

Exercise 9.6.2.

Solution. Since $T \wedge n$ is a bounded stopping time, by optional stopping theorem we have $\mathbb{E}(S^2_{T\wedge n} - T \wedge n) = \mathbb{E}(S_0^2 - 0) = 0$. So $\mathbb{E}(S^2_{T\wedge n}) = \mathbb{E}(T \wedge n)$. By DCT and MCT, $\lim_{n\to\infty} \mathbb{E}(S^2_{T\wedge n}) = \mathbb{E}(S_T^2) = \mathbb{E}(T) = \lim_{n\to\infty} \mathbb{E}(T \wedge n)$. Hence

$$\mathbb{E}(T) = \mathbb{E}(S_T^2) = a^2 \mathbb{P}(S_T = a) + b^2 \mathbb{P}(S_T = b) = \frac{a^2 b}{b-a} - \frac{b^2 a}{b-a}.$$

□

Exercise 9.6.7.

Solution. Since $\mathbb{E}(S_n - \mu n \mid \mathcal{F}_{n-1}) = \mathbb{E}X_n + S_{n-1} - \mu n = S_{n-1} - \mu(n-1)$, $S_n - \mu n$ is a martingale adapted to $\{\mathcal{F}_n\}$. Then by optional stopping theorem and $T \wedge n$ is a bounded stopping time, $\mathbb{E}\big(S_{T\wedge n} - \mu(T \wedge n)\big) = \mathbb{E}(S_0 - \mu \cdot 0) = 0$. Now by DCT and MCT we have $\mathbb{E}(S_T) = \mu\mathbb{E}(T)$.

9.7.

Exercise 9.7.4.

Solution. For each point in $\{T > n\}$, X_n is an odd number which is not a prime. And $X_n \leq p_{n-1} + 2$

where p_{n-1} is the largest prime divisor of X_{n-1}. That is,

$$X_n \leq \frac{X_{n-1}}{3} + 2 \leq \frac{X_{n-1}}{\sqrt{2}} \text{ a.s. on } \{T > n\}.$$

By monotonicity of conditional expectation and $\{T > n\} \in \mathcal{F}_{n-1}$,

$$\mathbb{E}(X_n \mathbb{1}_{\{T>n\}} | \mathcal{F}_{n-1}) \leq \mathbb{E}\Big(\frac{X_{n-1}}{\sqrt{2}} \mathbb{1}_{\{T>n\}} \Big| \mathcal{F}_{n-1}\Big) \leq \frac{X_{n-1}}{\sqrt{2}} \mathbb{1}_{\{T>n\}}.$$

Hence on $\{T > n\}$ a.s., by Jenson's inequality for concave function,

$$\mathbb{E}(\log(X_n)|\mathcal{F}_n) \leq \log(X_{n-1}) - \frac{1}{2}\log(2).$$

Since $\log(X_{T\wedge n}) \geq 0$, by the conclusion in (9.7.2) we have

$$\mathbb{E}(T) \leq \frac{\log(x) - 0}{\frac{1}{2}\log(2)} = C\log(x).$$

□

Solution. Write $S_n = X_1 + \cdots + X_n$. For any w_m (they're measurable),

$$\mathbb{E}w_m(x + X)\big|_{x=S_n} = \mathbb{E}_{X_{n+1}} w_m(S_n + X_{n+1}) = \mathbb{E}\big(w_m(S_{n+1})|\mathcal{F}_n\big),$$

by the independence of S_n and X_{n+1}. So

$$\begin{aligned} w_{N-n}(S_n) &= (S_n - K)^+ \vee \mathbb{E}w_{N-n-1}(x + X)\big|_{x=S_n} \\ &= (S_n - K)^+ \vee \mathbb{E}_{X_{n+1}} w_{N-(n+1)}(S_{n+1}) \\ &= (S_n - K)^+ \vee \mathbb{E}\big(w_{N-(n+1)}(S_{n+1})|\mathcal{F}_n\big), \end{aligned}$$

which means $w_{N-n}(S_n) \geq (S_n - K)^+$ and $\mathbb{E}\big(w_{N-(n+1)}(S_{n+1})|\mathcal{F}_n\big) \leq w_{N-n}(S_n)$, therefore $w_{N-n}(S_n)$ is a Snell envelope. Hence the optimal τ would be

$$\tau = \min\{n : (S_n - K)^+ = w_{N-n}(S_n)\}.$$

$w_{N-n}(S_n)$ represents the optimal payoff estimated at time n. At time $n-1$, to estimate the optimal payoff, we compare the current payoff and the average of optimal payoff estimated at time n.

□

For non-negative supermartingale,

$$(a - b)\mathbb{E}(U_m) \geq \sum_{n=1}^{m} \mathbb{E}(Z_n(Y_n - Y_{n-1})) \geq \sum_{n=1}^{m} \mathbb{E}(Y_n - Y_{n-1})$$

9.9

$\{n-1, n\}$-in-upcrossing indicator Z_n:

1. $Z_n = 1$: $n-1$ is in an upcrossing and $Y_{n-1} < b$.
2. $Z_n = 0$: Either (a) $n-1$ is in an upcrossing and $Y_{n-1} \geq b$, or (b) $n-1$ is not in an upcrossing.

So Z_n is $\mathcal{F}_{n-1}$-measurable.

9.11

Exercise 9.11.1. Let $A \in \mathcal{E}$. Then since $A \in \mathcal{G} = \sigma(X_1, X_2, \cdots)$, there exists some $B \subset \mathbb{R}^{\mathbb{N}}$ such that $A = \{(X_i)_{i\geq 1} \in B\}$. Given $\epsilon > 0$, since the algebra $\cup_{i\geq 1}\sigma(X_1, \cdots, X_i)$ generates $\mathcal{G}$, by the **approximation lemma** we know that there exists some $n \in \mathbb{N}$ and $A_n \in \sigma(X_1, \cdots, X_n)$ such that $\mathbb{P}(A\Delta A_n) < \epsilon$. Note that we can also find some $B_n \subset \mathbb{R}^n$ such that $A_n = \{(X_1, \cdots, X_n) \in B_n\}$.

Next, define random vector $(Y_i)_{i\geq 1} = h_n((X_i)_{i\geq 1})$, where

$$h_n : (x_1, x_2, \cdots) \mapsto (x_{n+1}, \cdots, x_{2n}, x_1, \cdots, x_n, x_{2n+1}, \cdots).$$

Then by the $i.i.d.$ assumption of r.v.s $\{X_i\}_{i\geq 1}$, $(Y_i)_{i\geq 1}$ and $(X_i)_{i\geq 1}$ have the same law, hence

$$\begin{aligned} \mathbb{P}(A\Delta A_n) &= \mathbb{P}\left(\{(X_i)_{i\geq 1} \in B\}\Delta\{(X_i)_{i=1}^n \in B_n\}\right) \\ &= \mathbb{P}\left(\{(Y_i)_{i\geq 1} \in B\}\Delta\{(Y_i)_{i=1}^n \in B_n\}\right) \\ (A'_n := \{(Y_i)_{i=1}^n \in B_n\}) &= \mathbb{P}(A\Delta A'_n) < \epsilon. \end{aligned}$$

The last equality holds because $A \in \mathcal{E}$ and h_n is a $2n$-permutation,

$$\{(Y_i)_{i\geq 1} \in B\} = \{(X_i)_{i\geq 1} \in B\} = A.$$

So we now have both $\mathbb{P}(A_n) \to \mathbb{P}(A)$ and $\mathbb{P}(A'_n) \to \mathbb{P}(A)$, and hence

$$\mathbb{P}(A_n)\mathbb{P}(A'_n) \to \mathbb{P}(A)^2.$$

Moreover,

$$\begin{aligned} \mathbb{P}(A_n\Delta A'_n) = \mathbb{E}\left|\mathbb{1}_{A_n} - \mathbb{1}_{A'_n}\right| &\leq \mathbb{E}\left|\mathbb{1}_{A_n} - \mathbb{1}_{A}\right| + \mathbb{E}\left|\mathbb{1}_{A} - \mathbb{1}_{A'_n}\right| \\ &= \mathbb{P}(A_n\Delta A) + \mathbb{P}(A\Delta A'_n) < 2\epsilon, \end{aligned}$$

which implies $\mathbb{P}(A_n) - \mathbb{P}(A_n \cap A'_n) < 2\epsilon$ and hence, $|\mathbb{P}(A) - \mathbb{P}(A_n \cap A'_n)| < 3\epsilon$. So we have

$$\mathbb{P}(A_n \cap A'_n) \to \mathbb{P}(A).$$

But $\{X_i\}_{i\geq 1}$ are independent, hence the events $A_n \in \sigma\left((X_i)_{i=1}^n\right)$ and $A'_n \in \sigma\left((Y_i)_{i=1}^n\right) = \sigma\left((X_i)_{i=n+1}^{2n}\right)$ are independent,

$$\mathbb{P}(A_n \cap A'_n) = \mathbb{P}(A_n)\mathbb{P}(A'_n) \to \mathbb{P}(A).$$

Therefore $\mathbb{P}(A)^2 = \mathbb{P}(A)$, implies $\mathbb{P}(A) = 0$ or 1.

□

Recall. Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For any $A \in \mathcal{G} = \sigma(X_1, X_2, \cdots)$ and any $\epsilon > 0$, show that there is some $n \geq 1$ and some $A_n \in \sigma(X_1, \cdots, X_n)$ such that $\mathbb{P}(A\Delta A_n) < \epsilon$.

Proof. We set $\mathcal{F} = \bigcup_{n\geq 1}\sigma(X_1, \cdots, X_n)$, and therefore $\mathcal{G} = \sigma(\mathcal{F})$. Next we check that $\mathcal{F}$ is an algebra.

1. Clearly $\Omega \in \mathcal{L}$.

2. Suppose $C \in \mathcal{F}$. Then $A \in \sigma(X_1, \cdots, X_n)$ for some $n \geq 1$, therefore $C^c \in \sigma(X_1, \cdots, X_n) \subset \mathcal{F}$.

3. Take any finite collection $A_1, \cdots, A_m \in \mathcal{F}$, then $A_1, \cdots, A_m \in \sigma(X_1, \cdots, X_M)$ for some $M \geq 1$ and therefore $\cup_{i\geq 1}^m A_i \in \mathcal{F}$.

Hence by **approximation lemma**, there exists $A_\epsilon \in \mathcal{F}$ such that $\mathbb{P}(A_\epsilon \Delta A) < \epsilon$.

Now since $\mathcal{F} \subseteq \mathcal{L}$, $\mathcal{G} = \sigma(\mathcal{F}) \subseteq \sigma(\mathcal{L})$. Therefore for any set $B \in \sigma(\mathcal{L})$ (also for any set in $\mathcal{G}$, of course), there exists $A \in \mathcal{L}$ such that $\mathbb{P}(A \Delta B) < \epsilon$.

$\square$

9.13.

Exercise 9.13.4. Suppose $\{X_n\}_{n\geq 0}$ is a martingale that is uniformly bounded by some $M > 0$ in L^2. Then for any $0 \leq n \leq m$, $\|X_m - X_n\|_2 \leq 2M$ and

$$\begin{aligned}\mathbb{E}(X_m - X_n)^2 &= \mathbb{E}\left[\mathbb{E}\left((X_m - X_n)^2|\mathcal{F}_n\right)\right]\\ &= \mathbb{E}X_m^2 - \mathbb{E}\left[2X_n\mathbb{E}(X_m|\mathcal{F}_n) - X_n^2\right]\\ &= \mathbb{E}X_m^2 - \mathbb{E}X_n^2.\end{aligned}$$

From this, we can write

$$\mathbb{E}(X_m - X_n)^2 = \sum_{i=n}^{m-1}\left(\mathbb{E}X_{i+1}^2 - \mathbb{E}X_i^2\right) = \sum_{i=n}^{m-1}\mathbb{E}(X_{i+1} - X_i)^2,$$

which shows that $\|X_n - X_0\|_2$ is a non-decreasing sequence of n. Therefore together with the uniformly bounded condition,

$$\lim_{n\to\infty}\|X_n - X_0\|_2^2 = \sum_{i=0}^{\infty}\mathbb{E}(X_{i+1} - X_i)^2 < \infty.$$

This means $\|X_m - X_n\|_2^2 = \sum_{i=n}^{m-1}\mathbb{E}(X_{i+1} - X_i)^2 < \epsilon$ for any sufficiently large n, m, and hence X_n is Cauchy in L^2. Then by the completeness of $\|\cdot\|_2$ in L^2 space, $X_n \to X$ in L^2 for some $X \in L^2$.

□

Exercise 9.13.6. Suppose $p > 1$. Since X_n is a martingale or a nonnegative submartingale, $|X_n|^p$ is a nonnegative submartingale. Suppose X_n is uniformly bounded in L^p. Then by *Theorem 9.13.3*, we have $\mathbb{E}|X_n - X|^p \to 0$ for some $X \in L_p$. Since $t \mapsto |t|^p$ is a convex function, for any $\lambda \in (0, 1)$,

$$|X_n|^p = \left|(1-\lambda)\frac{X_n - X}{1-\lambda} + \lambda\frac{X}{\lambda}\right|^p \leq (1-\lambda)^{1-p}|X_n - X|^p + \lambda^{1-p}|X|^p. \tag{1}$$

Given any $\epsilon > 0$, take $\lambda = \left(1 + \frac{\epsilon}{\|X\|_p^p}\right)^{1/(1-p)} \in (0, 1)$, then by (1) we have

$$\begin{aligned}\left||X_n|^p - |X|^p\right| &\leq \epsilon\cdot\frac{|X|^p}{\|X\|_p^p} + \left[1 - \left(\frac{\|X\|_p^p}{\|X\|_p^p + \epsilon}\right)^{\frac{1}{p-1}}\right]|X_n - X|^p\\ &< \epsilon\cdot\frac{|X|^p}{\|X\|_p^p} + |X_n - X|^p.\end{aligned}$$

Take expectation on both sides,

$$\mathbb{E}\left||X_n|^p - |X|^p\right| \leq \epsilon + \|X_n - X\|_p^p.$$

By the L^p convergence of X_n, $\|X_n - X\|_p^p \to 0$ as $n \to \infty$. Thus

$$\lim_{n\to\infty}\||X_n|^p - |X|^p\|_1 \leq \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have shown that $|X_n|^p$ converges in L^1 to $|X|^p$. Then it follows by *Proposition 8.3.5* that $|X_n|^p$ is uniformly integrable.

□

Exercise 9.13.7. Write the process $\{X_n^2\}$ as

$$X_n^2 = X_n^2 - \langle X\rangle_n + \langle X\rangle_n$$

where $\{X_n^2 - \langle X\rangle_n\}_{n\geq 0}$ is a martingale, since

$$\dots$$

Then by Doob's inequality for non-negative sub-martingale,

$$\mathbb{E}\max_{i\leq n}(X_i^2) \leq 4\mathbb{E}X_n^2 = 4\mathbb{E}\langle X\rangle_n.$$

Apply MCT on both sides, then we obtained

$$\mathbb{E}\sup_{n\geq 0} X_n^2 \leq 4\mathbb{E}\lim_{n\to\infty}\langle X\rangle_n = 4\mathbb{E}\langle X\rangle_\infty. \tag{2}$$

To discuss the property on the set $\{\langle X\rangle_\infty < \infty\}$, we start it by looking at $\{\langle X\rangle_\infty \leq k\}$, where $k \in \mathbb{N}$. Consider a random time

$$T_k := \max\{n : \langle X\rangle_n \leq k\}.$$

Then T_k is a stopping time, since $\langle X\rangle_n$ is non-decreasing and predictable:

$$\{T_k = n\} = \{\langle X\rangle_n \leq k\} \cap \{\langle X\rangle_{n+1} > k\} \in \mathcal{F}_n.$$

Therefore $X_{n\wedge T_k}$ is also a submartingale. By (2),

$$\sup_{n\geq 0}\mathbb{E}X_{n\wedge T_k}^2 \leq \mathbb{E}\sup_{n\geq 0} X_{n\wedge T_k}^2 \leq 4k.$$

Hence, by the submartingale convergence theorem (Theorem 9.13.3), $\{X_{n\wedge T_k}\}$ converges a.s. and in L^2 to some r.v. Notice that $X_{n\wedge T_k} = X_n$ on $\{T_k = \infty\}$. But $\{\langle X\rangle_\infty \leq k\} = \{T_k = \infty\}$, so X_n converges a.s. on $\{\langle X\rangle_\infty \leq k\}$. Since $k \in \mathbb{N}$ is arbitrary, we have X_n converges a.s. on $\{\langle X\rangle_\infty < \infty\}$.

□

Exercise 9.13.8 (Galton–Watson branching process).

(2) Since M_n is a martingale, $\mathbb{E}M_n = \mathbb{E}M_1 = \mathbb{E}X_{0,1} < \infty$ for all $n \geq 1$, therefore by the submartingale convergence theorem (*Theorem 9.9.2*), $M_n \to M$ a.s. and $\mathbb{E}|M| < \infty$. In particular, if $\mu < 1$, then $Z_n/\mu^n \to 0$ implies $Z_n \to 0$ as $n \to \infty$.

(3) Suppose $\mu = 1$. Then $Z_n = M_n$, and hence Z_n converges a.s. by (2). Also, we set $p_0 =$

$\mathbb{P}(X_{n,i}=0)$. Let $k, N \in \mathbb{N}$. Then by the definition of Z_n, we can write

$$\mathbb{P}\left(\bigcap_{n\geq N}\{Z_n=k\}\right)=\mathbb{P}\left(\{Z_N=k\}\cap\bigcap_{n>N}\left\{\sum_{i=1}^{k}X_{n-1,i}=k\right\}\right). \tag{3}$$

Note that Z_N is $\sigma(X_{0,1}, X_{1,1}, X_{1,2}, \cdots, X_{N-1,1}, X_{N-1,2}, \cdots)$-measurable, and the n-th summation set is $\sigma(X_{n-1,1}, X_{n-1,2}, \cdots)$-measurable. So by the i.i.d. assumption of $X_{n,i}$, the RHS is

$$\begin{aligned}\mathbb{P}\left(\{Z_N=k\}\cap\bigcap_{n>N}\left\{\sum_{i=1}^{k}X_{n-1,i}=k\right\}\right)&=\mathbb{P}(Z_N=k)\cdot\prod_{n>N}\mathbb{P}\left(\sum_{i=1}^{k}X_{n-1,i}=k\right)\\&=\mathbb{P}(Z_n=k)\cdot\prod_{n>N}\mathbb{P}\left(\sum_{i=1}^{k}X_{1,i}=k\right).\end{aligned}$$

But since

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{k}X_{1,i}=k\right)&\leq 1-\mathbb{P}(X_{1,1}=0, X_{1,2}=0, \cdots, X_{1,k}=0)\\&=1-\prod_{i=1}^{k}\mathbb{P}(X_{1,1}=0)=1-p_0^k<1,\end{aligned}$$

we find that the RHS of (3) is zero, therefore $\mathbb{P}\left(\bigcap_{n\geq N}\{Z_n=k\}\right)=0$. Taking the union bounds, we then see that

$$\mathbb{P}\left(\lim_{n\to\infty}Z_n=k\right)=\mathbb{P}\left(\bigcup_{N\geq 1}\bigcap_{n\geq N}\{Z_n=k\}\right)\leq\sum_{N\geq 1}\mathbb{P}\left(\bigcap_{n\geq N}\{Z_n=k\}\right)=0.$$

Since $k \in \mathbb{N}$ is arbitrary, we've proven that $\mathbb{P}\left(\lim_{n\to\infty}Z_n=0\right)=1$.

□

(4)

(5)

(6) By Theorem 9.13.3, $M_n \to M$ a.s. and in L^2, thus $M_n \to M$ in L^1 by Jenson's inequality. By Fatou's lemma,

$$\mathbb{E}M\leq\liminf_{n\to\infty}\mathbb{E}M_n\leq\limsup_{n\to\infty}\mathbb{E}M_n\leq\mathbb{E}M+\limsup_{n\to\infty}\mathbb{E}|M_n-M|=\mathbb{E}M.$$

Therefore $\mathbb{E}M=\lim_{n\to\infty}\mathbb{E}M_n=\mathbb{E}M_1>0$, so $\mathbb{P}(M>0)>0$.

□

Exercise 9.13.9. Suppose $X \in L^p$ for some $p \geq 1$. Recall the setting in **Exercise 9.12.2** that with the filtration $\mathcal{F}_n=\sigma\left(\cup_{i=0}^{2^n-1}\left[\frac{i}{2^n},\frac{i+1}{2^n}\right)\right)$, the process $X_n := \mathbb{E}(X|\mathcal{F}_n)$ is a martingale. Since $\mathbb{E}|X_n|^p=\mathbb{E}|X|^p<\infty$ for all $n\geq 1$, by the martingale L^p convergence theorem (*Theorem 9.13.3*) we can know that $X_n \to X$ in L^p.

$\square$

Exercise 9.13.11.

(2) **(Uniformly bounded)** Let $\phi_N(x) := \sqrt{N}(\cos \pi x)^{2N}$ for $N \geq 1$. Then for $0 < a < b < 1$, $x \in [0,1]$,

$$f_N(x) = \int_{a-x}^{b-x} \phi_N(t)dt.$$

Claim 1. For any $N \geq 1$ and $x \in [0,1]$,

$$f_N(x) \leq \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} \phi_N(t)dt = f_N\left(\frac{a+b}{2}\right).$$

Proof. Since $\phi_n(t) = \sqrt{n}(\cos \pi t)^{2n}$ is 1-periodic, $f_n(x) := \int_{a-x}^{b-x} \phi_n(t)dt$ is also 1-periodic. The derivative of f_n is

$$\frac{d}{dx} f_n(x) = -\phi_n(b-x) + \phi_n(a-x),$$

and we can easily see that

$$f_n'(x) = 0 \Leftrightarrow \frac{b+a}{2} - x = \frac{k}{2},\, k \in \mathbb{Z}$$

by the property of ϕ_n. Also, by the 1-periodicity, the global maxima must be $f(\frac{b+a}{2} + k)$ or $f(\frac{b+a+1}{2} + k)$, $k \in \mathbb{Z}$. If $\Delta := b - a \leq \frac{1}{2}$, since $\phi_n(s) \geq \phi_n(t)$ for all $|s| \leq \frac{\Delta}{2}$ and $|t - \frac{1}{2}| \leq \frac{\Delta}{2}$,

$$f\left(\frac{b+a-1}{2}\right) = \int_{\frac{1}{2}-\frac{\Delta}{2}}^{\frac{1}{2}+\frac{\Delta}{2}} \phi_n \leq \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \phi_n = f\left(\frac{b+a}{2}\right).$$

If $\Delta > \frac{1}{2}$,

$$\begin{aligned} f\left(\frac{b+a}{2}\right) = 2\int_0^{\frac{\Delta}{2}} \phi_n &= 2\left(\int_{\frac{1}{2}-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \phi_n + \int_0^{\frac{1}{2}-\frac{\Delta}{2}} \phi_n\right) \\ &\geq 2\left(\int_{\frac{1}{2}-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \phi_n + \int_{\frac{\Delta}{2}}^{\frac{1}{2}} \phi_n\right) = f\left(\frac{b+a-1}{2}\right). \end{aligned}$$

Therefore, $f(\frac{b+a}{2} + k)$, $k \in \mathbb{Z}$ is the maximum. Hence the claim follows. $\square$

We keep using the notation $\Delta = b - a$ in the following, and abbreviate $f_n\left(\frac{a+b}{2}\right)$ as f_n. First, write

$$\begin{aligned}
&\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}}(\cos\pi t)^{2n}dt = \frac{1}{\pi}(\cos\pi t)^{2n-1}\sin\pi t\Big|_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} + (2n-1)\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}}(\cos\pi t)^{2n-2}(\sin\pi t)^2 dt \\
&\Rightarrow 2n\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}}(\cos\pi t)^{2n}dt = \frac{2}{\pi}\left(\cos\frac{\pi\Delta}{2}\right)^{2n-1}\sin\frac{\pi\Delta}{2} + (2n-1)\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}}(\cos\pi t)^{2n-2}dt \\
&\Rightarrow f_n = \underbrace{\frac{1}{\pi\sqrt{n}}\left(\cos\frac{\pi\Delta}{2}\right)^{2n-1}\sin\frac{\pi\Delta}{2}}_{b_n} + \underbrace{\frac{2n-1}{2\sqrt{n(n-1)}}}_{c_n} f_{n-1} = b_n + c_n f_{n-1}.
\end{aligned}$$

Since $c_n > 1$ and $b_n > 0$ for all $n \geq 1$ (since $0 < \Delta < 1$), we can see that $\{f_n\}$ is a nonnegative monotonically increasing sequence. Moreover,

$$f_n = b_n + c_n b_{n-1} + c_n c_{n-1} b_{n-2} + \cdots + (c_n c_{n-1}\cdots c_3)b_2 + (c_n c_{n-1}\cdots c_2)f_1. \tag{4}$$

Claim 2. The product $\prod_{k=n}^{m} c_k$ is bounded above and below by some constants for all $2 \leq n \leq m$.

Proof. Clearly 1 is an lower bound since $c_n > 1$ for all $n \geq 1$. For the upper bound, we evaluate

$$\prod_{i=2}^{n} c_i = \prod_{i=2}^{n}\frac{2i-1}{2\sqrt{i(i-1)}} = \prod_{i=2}^{n}\sqrt{\frac{i}{i-1}}\frac{2i-1}{2i} = \sqrt{n}\prod_{i=2}^{n}\frac{2i-1}{2i}.$$

Note that the sequence $(1-\frac{1}{2n}) \uparrow 1$ as $n \to \infty$, so we have the following bound:

$$\log\prod_{i=2}^{n}\left(1-\frac{1}{2i}\right) = \sum_{i=2}^{n}\log\left(1-\frac{1}{2i}\right) \leq \int_1^{n+1}\log\left(1-\frac{1}{2x}\right)dx.$$

The right hand side is calculated as

$$\begin{aligned}
\int_1^{n+1}\log\left(1-\frac{1}{2x}\right)dx &= \int_1^{n+1}\log(2x-1) - \log(2x)dx \\
&= \frac{1}{2}\left[(u\log u - u)\Big|_1^{2n+1} - (u\log u - u)\Big|_2^{2n+2}\right] \\
&= \frac{1}{2}\left(\log\frac{(2n+1)^{2n+1}}{(2n+2)^{2n+2}} + 2\log 2\right) \\
&= \log\sqrt{\frac{4(2n+1)^{2n+1}}{(2n+2)^{2n+2}}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sqrt{n}\prod_{i=2}^{n}\frac{2i-1}{2i} &\leq \sqrt{\frac{4n(2n+1)^{2n+1}}{(2n+2)^{2n+2}}} = \sqrt{\frac{4n}{2n+1}}\sqrt{\left(1-\frac{1}{2n+2}\right)^{2n+2}} \\
&= \sqrt{2\left(1-\frac{1}{2n+1}\right)}\sqrt{\left(1-\frac{1}{2n+2}\right)^{2n+2}} \uparrow \frac{2}{\sqrt{e}}
\end{aligned} \tag{5}$$

as $n \to \infty$. Hence $\frac{2}{\sqrt{e}}$ is an upper bound. $\square$

Now by (4) and (5) we have

$$f_n \leq \frac{2}{\sqrt{e}} \left(\sum_{i=2}^{n} b_i + f_1 \right),$$

where

$$\sum_{i=2}^{n} b_i = \sum_{i=2}^{n} \frac{1}{\pi\sqrt{i}} \left(\cos \frac{\pi\Delta}{2} \right)^{2i-1} \sin \frac{\pi\Delta}{2} = \frac{1}{\pi} \tan \frac{\pi\Delta}{2} \sum_{i=2}^{n} \frac{1}{\sqrt{i}} \left(\cos \frac{\pi\Delta}{2} \right)^{2i}.$$

Notice that the sequence $\frac{1}{\sqrt{n}} \left(\cos \frac{\pi\Delta}{2}\right)^{2n} \downarrow 0$ as $n \to \infty$, hence we can write

$$\begin{aligned}
\sum_{i=2}^{n} \frac{1}{\sqrt{i}} \left(\cos \frac{\pi\Delta}{2} \right)^{2i} &\leq \int_1^n \frac{1}{\sqrt{x}} \left(\cos \frac{\pi\Delta}{2} \right)^{2x} dx \\
&= \int_1^n x^{-\frac{1}{2}} \exp \left[2x \log \left(\cos \frac{\pi\Delta}{2} \right) \right] dx \\
&\leq \frac{\Gamma\left(\frac{1}{2}\right)}{\sqrt{2 \log \left(\frac{1}{\cos \frac{\pi\Delta}{2}} \right)}}.
\end{aligned}$$

Where the last inequality is obtained by bounding via integrating on $(0, \infty)$. Finally, the upper bound of f_n is

$$f_n \left(\frac{a+b}{2} \right) \leq \frac{2}{\sqrt{e}} \left[\frac{\Gamma\left(\frac{1}{2}\right) \tan \frac{\pi\Delta}{2}}{\pi \sqrt{2 \log \left(\frac{1}{\cos \frac{\pi\Delta}{2}} \right)}} + \frac{1}{2} \left(\Delta + \frac{1}{\pi} \sin \pi\Delta \right) \right] < \infty$$

for all $n \geq 1$. Since we've already know that $f_n(\frac{a+b}{2})$ is increasing, $f_n\left(\frac{a+b}{2}\right) \uparrow L$ for some $L < \infty$. Therefore, together with *Claim 1*, $f_n(x)$ is bounded by L for all $n \geq 1$ on $x \in [0, 1]$.

(Convergence)

$\square$

(3) By (2) we can see that $|g(x)f_N(x)| \leq c|g(x)|$ for all $N \geq 1$ on $[0, 1]$, hence by dominated convergence theorem,

$$\lim_{N\to\infty} \int g(x) f_N(x) dx = \int c g(x) \mathbb{1}_{(a,b)} dx = c \int_a^b g(x) dx. \tag{6}$$

Where $c\mathbb{1}_{(a,b)}(x) = \lim_{N\to\infty} f_N(x)$ a.e. on $x \in [0,1]$. Now fix a $N \geq 1$,

$$\begin{aligned}\int_0^1 g(x) f_N(x)dx &= \int_0^1 \int_a^b g(x)\sqrt{N}\left(\frac{1+\cos 2\pi(u-x)}{2}\right)^N dudx \\ &= \int_0^1 \int_a^b g(x)\sqrt{N}\left(\cos\pi(u-x)\right)^{2N} dudx \\ &= \int_a^b \sqrt{N}\int_0^1 g(x)(\cos\pi(u-x))^{2N} dxdu,\end{aligned} \tag{7}$$

where we expand $(\cos\pi(u-x))^{2N}$ as

$$\begin{aligned}(\cos\pi(u-x))^{2N} &= \left(\frac{e^{i\pi(u-x)}+e^{-i\pi(u-x)}}{2}\right)^{2N} \\ &= \frac{e^{-i2\pi N(u-x)}}{4^N}(e^{i2\pi(u-x)}+1)^{2N} \\ &= \frac{e^{-i2\pi N(u-x)}}{4^N}\sum_{k=0}^{2N}\binom{2N}{k}e^{i2\pi k(u-x)} \\ &= \frac{1}{4^N}\sum_{k=-N}^{N}\binom{2N}{k+N}e^{i2\pi k(u-x)} \\ &= \frac{1}{4^N}\left[\binom{2N}{N}+2\sum_{k=1}^{N}\binom{2N}{k+N}\cos 2\pi k(u-x)\right].\end{aligned}$$

Also note that $\cos 2\pi k(u-x) = \cos 2\pi ku\cos 2\pi kx + \sin 2\pi ku \sin 2\pi kx$. This implies that the integral

$$\int_0^1 g(x)(\cos\pi(u-x))^{2N}dx = 0,$$

since $(\cos\pi(u-x))^{2N}$ is a linear combination of $\{1, \cos 2\pi nx, \sin 2\pi nx\}_{n\geq 1}$ and g is orthogonal to them. Therefore by (7), $\int_0^1 g(x)f_N(x)dx = 0$. Thus by (6),

$$\int_a^b g(x)dx = 0.$$

Since $0 < a < b < 1$ is arbitrary, we can know that if g_n is defined on $[0,1)$ as

$$g_n(x) = \sum_{i=0}^{2^n-1} 2^n\left[\int_{\frac{i}{2^n}}^{\frac{i+1}{2^n}} g(u)du\right]\mathbb{1}_{[\frac{i}{2^n},\frac{i+1}{2^n})}(x),$$

then we have $g_n(x) = 0$ on $[0,1)$. Also, **Exercise 9.12.2** has shown that $g_n \to g$ a.e. as $n\to\infty$. Therefore, $g = 0$ a.e. on $[0,1]$.

$\square$

9.14

Exercise 9.14.3.

(2) It's clear that for any $n \geq 0$,

$$\int_{a_n}^{a_{n+1}} \frac{1}{x^2} dx \geq \int_{a_n}^{a_{n+1}} \frac{1}{a_{n+1}^2} dx.$$

Therefore

$$\sum_{n\geq 0} \frac{a_{n+1}-a_n}{a_{n+1}^2} \leq \sum_{n\geq 0} \int_{a_n}^{a_{n+1}} \frac{1}{x^2} dx = \lim_{n\to\infty} \int_{a_0}^{a_n} \frac{1}{x^2} dx \leq \frac{1}{a_0} < \infty.$$

(1) First of all, if $a_n \uparrow A < \infty$, then similar to the above,

$$\sum_{n\geq 0} \frac{a_{n+1}-a_n}{a_n} \leq \lim_{n\to\infty} \int_{a_0}^{a_n} \frac{1}{x} dx = A - \log(a_0) < \infty.$$

If $a_n \uparrow \infty$, we discuss under two cases. Suppose $\liminf a_n/a_{n+1} = 0$. Then there is a subsequence $\{a_{n_k}\}_{k\geq 1}$ such that $a_{n_k}/a_{n_k+1} \to 0$ as $k \to \infty$. Hence for an $\epsilon \in (0,1)$, there exists an $N \geq 1$ such that

$$\sum_{n\geq 0} \left(1 - \frac{a_n}{a_{n+1}}\right) \geq \sum_{k>N} \left(1 - \frac{a_{n_k}}{a_{n_k+1}}\right) \geq \sum_{k>N} (1-\epsilon) = \infty.$$

Suppose $\liminf a_n/a_{n+1} = c > 0$. Then we can take an $N \geq 1$ such that

$$\frac{c}{2} < \inf_{n>N} \left(\frac{a_n}{a_{n+1}}\right) \leq c.$$

Define a number M as

$$M := \min_{0\leq n\leq N} \left(\frac{a_n}{a_{n+1}}\right) \wedge \frac{c}{2},$$

then $0 < M \leq \frac{a_n}{a_{n+1}}$ for all $n \geq 0$. Therefore

$$\begin{aligned}
\sum_{n\geq 0} \frac{a_{n+1}-a_n}{a_{n+1}} &\geq M \sum_{n\geq 0} \frac{a_{n+1}-a_n}{a_n} \\
&\geq M \sum_{n\geq 0} \int_{a_n}^{a_{n+1}} \frac{1}{x} dx = M \left[\lim_{n\to\infty} \log(a_n) - \log(a_0)\right] = \infty.
\end{aligned}$$

□

Exercise 9.14.4. Let $Z_n = S_n^2/c_n^2$. Since c_n is $\mathcal{F}_{n-1}$-measurable, and S_n is square-integrable martingale,

$$\begin{aligned}
\mathbb{E}(Z_n|\mathcal{F}_{n-1}) &= \frac{1}{c_n^2}\mathbb{E}\left(S_{n-1}^2 + S_n^2 - S_{n-1}^2|\mathcal{F}_{n-1}\right) \\
&= \frac{S_{n-1}^2}{c_{n-1}^2} + \frac{\mathbb{E}\left((S_n - S_{n-1})^2|\mathcal{F}_{n-1}\right)}{c_n^2} - \left(1 - \frac{c_{n-1}^2}{c_n^2}\right)\frac{S_{n-1}^2}{c_{n-1}^2} \\
&= Z_{n-1} + \frac{\sigma_n^2}{c_n^2} - \left(1 - \frac{c_{n-1}^2}{c_n^2}\right)\frac{S_{n-1}^2}{c_{n-1}^2}.
\end{aligned}$$

Also, since $c_{n-1} \le c_n$ for all $n \ge 1$, the last two terms are nonnegative and $\mathcal{F}_{n-1}$-measurable. Hence Z_n is a nonnegative almost supermartingale. By Robbins-Siegmund's theorem, we know that $Z_n \to Z$ a.s. for some Z that is finite a.s. and

$$\sum_{n=1}^{\infty}\left(\frac{c_n^2 - c_{n-1}^2}{c_n^2}\right)\frac{S_{n-1}^2}{c_{n-1}^2} < \infty$$

on the set $\{\sum_{n=1}^{\infty}\sigma_n^2/c_n^2 < \infty\}$. Moreover, since c_n^2 is positive and increasing, by **Exercise 9.14.3** we know that

$$\sum_{n=1}^{\infty}\frac{c_n^2 - c_{n-1}^2}{c_n^2} = \infty.$$

Hence if $\sum_{n=1}^{\infty}\sigma_n^2/c_n^2 < \infty$, we must have $S_n^2/c_n^2 \to 0$, and thus $S_n/c_n \to 0$ as $n \to \infty$.

□

Exercise 9.14.5. Let $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$ be the filtration. Suppose $S_n = \sum_{i=1}^{n} X_i X_{i-1}$. Then S_n is a $\mathcal{F}_n$-adapted martingale, since by the i.i.d. and zero-mean condition of X_n,

$$\mathbb{E}(S_n|\mathcal{F}_{n-1}) = S_{n-1} + \mathbb{E}(X_n)X_{n-1} = S_{n-1}.$$

Also,

$$\sigma_n^2 = Var(S_n|\mathcal{F}_{n-1}) = \mathbb{E}\left((S_n - S_{n-1})^2|\mathcal{F}_{n-1}\right) = \mathbb{E}(X_n^2 X_{n-1}^2|\mathcal{F}_{n-1}) = \sigma^2 X_{n-1}^2.$$

Here $\sigma^2 = \mathbb{E}X_0^2$ is a fixed constant. Then, by MCT,

$$\mathbb{E}\sum_{n\ge1}\frac{\sigma_n^2}{n^2} = \sigma^2\mathbb{E}\sum_{n\ge1}\frac{X_{n-1}^2}{n^2} = \sigma^2\sum_{n\ge1}\frac{\mathbb{E}X_{n-1}^2}{n^2} = \sum_{n\ge1}\frac{\sigma^4}{n^2} < \infty,$$

then we have $\sum_{n\ge1}\sigma_n^2/n^2 < \infty$ a.s. Hence by the previous exercise, we know that $S_n/n \to 0$ a.s.

□

Exercise 9.14.6. Define the random variable $P_n := \sum_{k=0}^{n} p_k = \sum_{k=0}^{n}\mathbb{P}(X_k = 1|\mathcal{F}_{k-1})$. We also define a sequence of nonnegative random variables $Z_n := (S_n - P_n)^2$. Then we consider the conditional expectation,

$$\mathbb{E}\left(\frac{Z_n}{P_n^2}\middle|\mathcal{F}_{n-1}\right) = \frac{1}{P_n^2}\mathbb{E}(Z_n - Z_{n-1}|\mathcal{F}_{n-1}) + \frac{Z_{n-1}}{P_{n-1}^2} - \left(1 - \frac{P_{n-1}^2}{P_n^2}\right)\frac{Z_{n-1}}{P_{n-1}^2}.$$

Rewrite the first term as

$$\begin{aligned}\mathbb{E}(Z_n - Z_{n-1}|\mathcal{F}_{n-1}) &= \mathbb{E}\left((S_n - P_n)^2|\mathcal{F}_{n-1}\right) - (S_{n-1} - P_{n-1})^2 \\ &= \mathbb{E}(S_n^2|\mathcal{F}_{n-1}) - 2P_n\mathbb{E}(S_n|\mathcal{F}_{n-1}) + P_n^2 - (S_{n-1} - P_{n-1})^2\end{aligned}$$

Notice that since $X_n = \mathbb{1}_{\{X_n=1\}} = (\mathbb{1}_{\{X_n=1\}})^2 = X_n^2$,

$$p_n = \mathbb{P}(X_n = 1|\mathcal{F}_{n-1}) = \mathbb{E}(X_n|\mathcal{F}_{n-1}) = \mathbb{E}(X_n^2|\mathcal{F}_{n-1}),$$

so we have

$$\begin{aligned}\mathbb{E}(S_n^2|\mathcal{F}_{n-1}) &= S_{n-1}^2 + 2p_nS_{n-1} + p_n \\ 2P_n\mathbb{E}(S_n|\mathcal{F}_{n-1}) &= 2P_n(S_{n-1} + p_n) \\ \mathbb{E}\left(P_n^2 - (S_{n-1} - P_{n-1})^2|\mathcal{F}_{n-1}\right) &= -S_{n-1}^2 + 2P_{n-1}S_{n-1} + 2P_{n-1}p_n + p_n^2\end{aligned}$$

Combine them together, we then get

$$\mathbb{E}(Z_n - Z_{n-1}|\mathcal{F}_{n-1}) = p_n - p_n^2 = p_n(1 - p_n) \leq p_n.$$

The last inequality holds since $0 \leq \mathbb{E}(\mathbb{1}_{\{X_n=1\}}|\mathcal{F}_{n-1}) \leq 1$. Therefore,

$$\mathbb{E}\left(\frac{Z_n}{P_n^2}\bigg|\mathcal{F}_{n-1}\right) \leq \frac{Z_{n-1}}{P_{n-1}^2} + \frac{p_n}{P_n^2} - \left(1 - \frac{P_{n-1}^2}{P_n^2}\right)\frac{Z_{n-1}}{P_{n-1}^2}.$$

Since Z_{n-1}/P_{n-1}^2, p_n/P_n^2 are nonnegative and $\mathcal{F}_{n-1}$-measurable, and $0 \leq P_{n-1} \leq P_n$, by Robbins-Siegmund's almost supermartingale theorem, we have $\left(\frac{S_n}{P_n} - 1\right)^2 \to L < \infty$ a.s. on $\left\{\sum_{n\geq 0}\frac{p_n}{P_n^2} < \infty\right\}$. Notice that by the first result of **Exercise 9.14.3**,

$$\sum_{n\geq 0}\frac{p_n}{P_n^2} = \sum_{n\geq 0}\frac{P_n - P_{n-1}}{P_n^2} < \infty \text{ a.s.}$$

And the second result,

$$\sum_{n\geq 0}\left(\frac{P_n^2 - P_{n-1}^2}{P_n^2}\right) = \infty,$$

We have $\left(\frac{S_n}{P_n} - 1\right)^2 \to 0$ a.s.

$\square$

10.2

10.2.5. $\varphi(x) := \sqrt{x}$ on $[0, 1]$ with Lebesgue measure λ.

$\square$

Exercise 10.2.6. (Measure preserving) Let $\omega = (\omega_e)_{e \in E} \in \{0, 1\}^E$. Since μ is the law of the i.i.d. Bernoulli(p) random variables attached to each $e \in E$, the law of $(\omega_e)_{e \in E}$ and $(\omega_e)_{e \in x+E}$ are the same, i.e. for all $A \subseteq \{0, 1\}^E$,

$$\mu((\omega_e)_{e \in E} \in A) = \mu((\omega_{x+e})_{e \in E} \in A).$$

Hence

$$\mu(T_x(\omega) \in A) = \mu((\omega_{x+e})_{e \in E} \in A) = \mu((\omega_e)_{e \in E} \in A).$$

(Ergodicity) Let $x \neq 0$. Let $(X_e)_{e \in E}$ be the random variables such that $X_e(\omega) = \omega_e$ for all $e \in E$. Then for any $k \geq 1$,

$$\{T_x^k \in A\} = \{(\omega_{kx+e})_{e \in E} \in A\} = \{(X_{kx+e}(\omega))_{e \in E} \in A\} \in \sigma((X_{kx+e})_{e \in E}).$$

Therefore

$$A \in \bigcap_{k \geq 1} \sigma((X_{kx+e})_{e \in E}) = \mathcal{T}.$$

Since $(X_e)_{e \in E}$ is i.i.d. Bernoulli(p) random variables under μ, by Kolmogorov's 0-1 law we have $\mu(A) \in \{0, 1\}$. Therefore T_x is an ergodic transformation.

$\square$

Exercise 10.2.7. Let $k \in \{0, 1, \cdots\} \cup \{\infty\}$. Then for $\omega \in \Omega$, $N(\omega) = k$ if and only if $N(T_x(\omega)) = k$, i.e.

$$\{N = k\} = \{T_k \in \{N = k\}\}.$$

Hence $\{N = k\} \in \mathcal{I}$. Since T_k is ergodic, $\mathcal{I}$ is trivial. So N is a constant a.s.

$\square$

Exercise 10.2.8. Suppose $N = k$ almost surely for some $k > 1$. Consider a d dimensional cube $B_N = \big\{(x_1, \cdots, x_d) \in \mathbb{Z}^d : -N \leq x_i \leq N\big\}$. Define

$$A_N := \{B_N \text{ intersects more than one } \infty \text{ cluster}\}.$$

Then

(1) $A_N \subseteq A_{N+1}$ for all $N \geq 1$.

(2) There exists some $N \geq 1$ s.t. $\omega \in A_N$ if and only if ω has more than one ∞ cluster.

Therefore

$$\lim_{N \to \infty} \mathbb{P}(A_N) = \mathbb{P}\left(\bigcup_{N \geq 1} A_N\right) = 1.$$

Hence there is some $N \geq 1$ such that $\mathbb{P}(A_N) > 0$. However, since A_N depends only on the edges outside B_N, A_N is independent of $\sigma\left((X_e)_{e\in E(B_N)}\right)$, so

$$\mathbb{P}(N_\infty = 1) \geq \mathbb{P}(A_N)p^{2(2N)(2N+1)} > 0,$$

which leads to a contradiction.

$\square$

Exercise 10.2.9. Define $u : \left[0, \frac{1}{2}\right] \to [0,1]$ as $u(x) = 4x(1-x)$ for all $0 \leq x \leq \frac{1}{2}$. Then u is a bijection defined on $\left[0, \frac{1}{2}\right]$, and its inverse is $u^{-1}(t) = \frac{1-\sqrt{1-t}}{2}$ for all $t \in [0,1]$.

$$\mathbb{P}(\varphi(x) \leq t) = \mathbb{P}\left(-4x(1-x) \geq -t\right) = \mathbb{P}\left(\left(x - \frac{1}{2}\right)^2 \geq \frac{1-t}{4}\right).$$

Since the density $\frac{1}{\pi\sqrt{x(1-x)}}$ is symmetric about $1/2$,

$$\mathbb{P}\left(\left(x - \frac{1}{2}\right)^2 \geq \frac{1-t}{4}\right) = 2\mathbb{P}\left(x \leq \frac{1}{2} - \sqrt{\frac{1-t}{4}}\right) = 2\mathbb{P}\left(x \leq u^{-1}(t)\right).$$

So

$$\frac{1}{2}\mathbb{P}\left(\varphi(x) \leq t\right) = \int_0^{u^{-1}(t)} \frac{1}{\pi\sqrt{x(1-x)}}dx.$$

Apply the change of variable $v = u(x)$ on the RHS, then it becomes

$$\int_0^{u(u^{-1}(t))} \frac{2}{\pi}\sqrt{\frac{1}{v}}\left|\frac{du^{-1}}{dv}\right| dv = \frac{1}{2\pi}\int_0^t \frac{1}{\sqrt{v(1-v)}}dv = \frac{1}{2}\mathbb{P}(x \leq t).$$

Therefore φ is a measure-preserving transform w.r.t. this $\mathbb{P}$. Since $\{[0,t) : 0 \leq t \leq 1\}$ generates $\mathcal{B}([0,1])$, ... So $\mathbb{P}(\varphi(x) \in A) = \mathbb{P}(x \in A)$ for all $A \in \mathcal{B}([0,1])$.

$\square$

Exercise 12.1.4.

Sol. (Separable) We claim that for any $f \in C[0,\infty)$, there is a sequence of functions of the following form,

$$p(x) = \sum_{j=0}^{\infty} p^{(j)}(x)\mathbb{1}_{[j,j+1]}(x),$$

where $p^{(j)}$ is the polynomial with rational coefficients, converge to f in ρ. Clearly, the set of functions of this form is countable since the cardinality is the same as $\mathbb{N}^2$ and $\mathbb{N}$.

Let $f \in C[0,\infty)$. By Weierstrass approximation theorem, for each $j \in \mathbb{Z}_{\geq 0}$, there is a sequence of polynomials with rational coefficients $(P_n^{(j)})_{n\geq 1}$ such that $\sup_{x\in[j,j+1]} |f(x) - P_n^{(j)}(x)| \to 0$ as $n \to \infty$. Define a sequence of functions

$$P_n(x) := \sum_{j=0}^{\infty} P_n^{(j)}(x)\mathbb{1}_{[j,j+1]}(x).$$

Then

$$\rho(f, P_n) = \sum_{j=0}^{\infty} 2^{-j} \frac{\|f - P_n\|_{[j,j+1]}}{1 + \|f - P_n\|_{[j,j+1]}} = \sum_{j=0}^{\infty} 2^{-j} \frac{\|f - P_n^{(j)}\|_{[j,j+1]}}{1 + \|f - P_n^{(j)}\|_{[j,j+1]}}.$$

Let $g_n : j \mapsto g_n(j) = \frac{\|f-P_n^{(j)}\|_{[j,j+1]}}{1+\|f-P_n^{(j)}\|_{[j,j+1]}}$ be a sequence of function on $\mathbb{Z}_{\geq 0}$. Then $|g_n| \leq 1$, and $g_n(j) \to 0$ as $n \to \infty$ since $\|f - P_n^{(j)}\|_{[j,j+1]} \to 0$ as $n \to \infty$. By the dominated convergence theorem,

$$\lim_{n\to\infty} \rho(f, P_n) = \lim_{n\to\infty} \sum_{j=0}^{\infty} 2^{-j} g_n(j) = \sum_{j=0}^{\infty} 2^{-j} \lim_{n\to\infty} g_n(j) = 0.$$

Therefore $(C[0,\infty), \rho)$ is separable.

(Complete) Let $(f_n)_{n\geq 0}$ be a Cauchy sequence in $(C[0,\infty), \rho)$. Fix $k \in \mathbb{Z}_{\geq 0}$. Choose n, m sufficiently large so that $\rho(f_n, f_m) < 2^{-k}\frac{\varepsilon}{1+\varepsilon}$, then

$$2^{-k} \frac{\|f_n - f_m\|_{[k,k+1]}}{1 + \|f_n - f_m\|_{[k,k+1]}} \leq \rho(f_n, f_m) < 2^{-k} \frac{\varepsilon}{1+\varepsilon},$$

and we get

$$\|f_n - f_m\|_{[k,k+1]} \leq \varepsilon.$$

Therefore $(f_n|_{[k,k+1]})_{n\geq 1}$ is a Cauchy sequence in $C[k, k+1]$ equipped with the sup-metric, which is a complete metric space by **Exercise 12.1.3**. Moreover, the sup-metric limit of $f_n|_{[k,k+1]}$ is its pointwise limit $\lim_{n\to\infty} f_n|_{[k,k+1]}(x)$. This ensures that $f_n(x)$ converges for all $x \in \mathbb{R}_{\geq 0}$. By setting $f : x \mapsto f(x) := \lim_{n\to\infty} f_n(x)$, we know that f_n converges to f uniformly on any interval of the type $[j, j+1]$. It follows that f_n converges uniformly to f on any compact sets in $\mathbb{R}_{\geq 0}$.

Also, f is indeed the ρ-limit of f_n: By dominated convergence,

$$\rho(f_n, f) \to \sum_{j\geq 0} 2^{-j} \lim_{n\to\infty} \frac{\|f - f_n\|_{[j,j+1]}}{1 + \|f - f_n\|_{[j,j+1]}} = 0.$$

Finally, since for any $n \in \mathbb{N}$, $\delta > 0$,

$$|f(x+\delta) - f(x)| \leq |f(x+\delta) - f_n(x+\delta)| + |f_n(x+\delta) - f_n(x)| + |f_n(x) - f(x)|.$$

Then, we bound the RHS by choosing a n to bound $|f_n(x+s) - f(x+s)|$ for any $s \in [-\Delta, \Delta]$ by the truth of uniform convergence, and then choose $\delta < \Delta$ to bound $|f_n(x) - f_n(x+\delta)|$. This proves that $f \in C[0, \infty)$, and therefore the space $(C[0, \infty), \rho)$ is complete.

Doob's optional stopping theorem:

1. The stopping time $\tau < \infty$ a.s.

2. $M_{t\wedge\tau}$ behaves regularly enough, e.g. $M_{t\wedge\tau}$ is uniformly bounded.

Then $\lim_{t\to\infty} M_{t\wedge\tau} = M_\tau$ a.s., and note that $M_{t\wedge\tau}$ is a martingale,

$$\mathbb{E} \lim_{t\to\infty} M_{t\wedge\tau} = \lim_{t\to\infty} \mathbb{E} M_{t\wedge\tau} = \lim_{t\to\infty} \mathbb{E} M_0 = \mathbb{E} M_0.$$

That is, $\mathbb{E} M_\tau = \mathbb{E} M_0$.