PTE 5TH EDITION SOLUTIONS MANUAL

HOIL LEE AND WONJUN SEO

Contents

Date: September 9, 2021.

1. Measure Theory

1.1. **Probability Theory.**

Exercise 1.1.1. (i) $\mathcal{F}$ is a σ-algebra.

From the definition, $\emptyset$, Ω are in $\mathcal{F}$, and $\mathcal{F}$ is closed under operation of complement. Now, suppose $A_1,\ A_2,\cdots$ are in $\mathcal{F}$. If all A_n's are countable, then $\cup_{n=1}^{\infty}A_n$, which is countable union of countable sets, is also countable, and thus is in $\mathcal{F}$. Otherwise, $\cup_{n=1}^{\infty}A_n^c$ is countable, and thus $\cup_{n=1}^{\infty}A_n$ is in $\mathcal{F}$.

(ii) P is a probability measure.

By definition, $P(A) \geq P(\emptyset) = 0$, for any $A \in \mathcal{F}$. For disjiont $A_1,\ A_2,\cdots \in \mathcal{F}$, if all A_n's are countable, then

$$P(\cup_{n=1}^{\infty}A_n) = 0 = \sum_{n=1}^{\infty} P(A_n).$$

If one of A_n is uncountable(say A_k), then since A_k^c is countable and $A_j \subset A_k^c$ for $j \neq k$, $(\cup_{n=1}^{\infty}A_n)^c \subset A_k^c$ and A_j's are countable for $j \neq k$. Thus,

$$P(\cup_{n=1}^{\infty}A_n) = 1 = P(A_k) + \sum_{n\neq k} A_n.$$

Therefore, $(\Omega, \mathcal{F}, P)$ is a probability space.

Exercise 1.1.2. Let $\mathcal{O}_d = \{(a_1, b_1) \times \cdots \times (a_d, b_d) : a_i, b_i \in \mathbb{R}\}$. It suffices to show that $\sigma(\mathcal{O}_d) = \sigma(\mathcal{S}_d)$. Since,

$$(a_1, b_1] \times \cdots \times (a_d, b_d] = \bigcap_{n=1}^{\infty} (a_1, b_1 + \frac{1}{n}) \times \cdots \times (a_d, b_d + \frac{1}{n}).$$

$$(a_1, b_1) \times \cdots \times (a_d, b_d) = \bigcup_{n=1}^{\infty} (a_1, b_1 - \frac{1}{n}) \times \cdots \times (a_d, b_d - \frac{1}{n})$$

Exercise 1.1.3. $\mathcal{C} = \{(a_1, b_1) \times \cdots \times (a_d, b_d) : \text{rational numbers } a < b\}$.

Exercise 1.1.4. Let $\mathcal{F} = \cup_i \mathcal{F}_i$.

(i) Clearly, $\emptyset$, $\Omega \in \mathcal{F}$. If $A \in \mathcal{F}$, then there exists j such that $A \in \mathcal{F}_j$. Then, $A^c \in \mathcal{F}_j$, and thus $A^c \in \mathcal{F}_j$. Let A and B are in $\mathcal{F}$. Then, there exist $j_1,\ j_2$ such that $A \in \mathcal{F}_{j_1}$ and $B \in \mathcal{F}_{j_2}$. WLOG, let $j_1 \leq j_2$. Since $\mathcal{F}_n$ are increasing sequence, A is also in $\mathcal{F}_{j_2}$. Thus, $A \cup B \in \mathcal{F}_j \subset \mathcal{F}$. Therefore, $\mathcal{F}$ is an algebra.

(ii) Let $\Omega = (0, 1]$ and $I_n = \{(\frac{j-1}{2^n}, \frac{j}{2^n}], \text{ for } j = 1, \ldots, 2^n\}$ for $^{\forall}n$. Define

$$\mathcal{F}_n \equiv \sigma(I_n) \text{ for } n = 1, 2, \cdots$$

Since $(\frac{j-1}{2^n}, \frac{j}{2^n}] = (\frac{2(j-1)}{2^{n+1}}, \frac{2j-1}{2^{n+1}}] \cup (\frac{2j-1}{2^{n+1}}, \frac{2j}{2^{n+1}}]$, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$.

Now,

$$(0, 1) = \bigcup_{n=1}^{\infty} (\frac{2^n - 2}{2^n}, \frac{2^n - 1}{2^n}],$$

where $(\frac{2^n-2}{2^n}, \frac{2^n-1}{2^n}] \in \mathcal{F}_n$ for $^\forall n$. However, since (0,1) $\notin \mathcal{F}_n$ for $^\forall n$, (0,1) $\notin \mathcal{F}$. Therefore, $\mathcal{F}$ is not a σ-algebra.

Exercise 1.1.5. (Counterexample) Let A be the set of even numbers. Then, $A \in \mathcal{A}$ with $\theta = \frac{1}{2}$. Let $I_0 = \{1\}, I_k = \{n : 2^{k-1} < n \leq 2^k\}$, for $k = 1, \cdots$. Construct B as follows.

$$B = \cup_{k=0}^{\infty}\{(I_{2k} \cap A^c) \cup (I_{2k+1} \cap A)\}$$

Then, $B \in \mathcal{A}$ with $\theta = \frac{1}{2}$. However, $A \cap B \notin \mathcal{A}$ for following reasons.

(1) $[n = 2^{2m}]$

$|(A \cap B) \cap \{1, 2, \ldots, 2^{2m}\}| = \sum_{k=1}^{m} \frac{1}{2}|I_{2k-1}|$.

$$\begin{aligned}
\frac{|(A \cap B) \cap \{1, 2, \ldots, 2^{2m}\}|}{2^{2m}} &= \frac{\sum_{k=1}^{m} \frac{1}{2}|I_{2k-1}|}{2^{2m}} \\
&= \frac{1 + \sum_{k=2}^{m} 2^{2k-3}}{2^{2m}} \\
&= \frac{\frac{2^{2m-1}+1}{3}}{2^{2m}} \to \frac{1}{6},
\end{aligned}$$

(2) $[n = 2^{2m+1}]$

Similarly,

$$\begin{aligned}
\frac{|(A \cap B) \cap \{1, 2, \ldots, 2^{2m+1}\}|}{2^{2m+1}} &= \frac{\sum_{k=1}^{m+1} \frac{1}{2}|I_{2k-1}|}{2^{2m+1}} \\
&= \frac{1 + \sum_{k=2}^{m+1} 2^{2k-3}}{2^{2m+1}} \\
&= \frac{\frac{2^{2m+1}+1}{3}}{2^{2m+1}} \to \frac{1}{3}
\end{aligned}$$

By (1), (2), $\limsup \frac{|(A\cap B)\cap\{1,2,\ldots,n\}|}{n} \geq \frac{1}{3}, \liminf \frac{|(A\cap B)\cap\{1,2,\ldots,n\}|}{n} \leq \frac{1}{6}$. Thus, $A \cap B \notin \mathcal{A}$. Therefore, $\mathcal{A}$ is neither an algebra nor a σ-algebra.

1.2. **Distributions.**

Exercise 1.2.1. To show that Z is a r.v., it suffices to show that for any $B \in \mathcal{B}(\mathbb{R})$, $Z^{-1}(B) \in \mathcal{F}$.

$$\begin{aligned}
Z^{-1}(B) &= [Z^{-1}(B) \cap A] \cup [Z^{-1}(B) \cap A^c] \\
&= [X^{-1}(B) \cap A] \cup [Y^{-1}(B) \cap A^c] \\
&\in \mathcal{B}(\mathbb{R}),
\end{aligned}$$

since X, Y are r.v.s, and $A \in \mathcal{F}$.

Exercise 1.2.2. Omit.

Exercise 1.2.3. Let A be the discontinuity set of distribution function F. Since F is right-continuous, for $x \in A$, $F(x^-) < F(x) = F(x^+)$ holds. For $x < y$ who are in A, we have

$F(x^-) < F(x) \leq F(y^-) < F(y)$. Pick rationals q_x and q_y such that $q_x \in (F(x^-), F(x))$ and $q_y \in (F(y^-), F(y))$. Then, $q_x < q_y$, and thus $|A| \leq \aleph_0$. Therefore, there exist at most countably many discontinuities.

Exercise 1.2.4. Let $F^{-1}(y) = \inf\{x : y \leq F(x)\}$. Then followings hold.

(i) $F(F^{-1}(y)) = y$ for $y \in (0,1)$.
(ii) F^{-1} is strictly increasing.
(iii) $\{F^{-1}(F(X)) \leq F^{-1}(y)\} = \{X \leq F^{-1}(y)\}$

(i) comes from continuity of F, and (ii) comes from monotonicity of F. For (iii),

$$\begin{aligned}\{F^{-1}(F(X)) \leq F^{-1}(y)\} &= \{F(X) \leq y\} && \text{(by (ii))}\\ &= \{X \leq F^{-1}(y)\}. && \text{(by definition)}\end{aligned}$$

Thus, for $0 < y < 1$,

$$\begin{aligned}P(F(X) \leq y) &= P(F^{-1}(F(X)) \leq F^{-1}(y)) && \text{(by (ii))}\\ &= P(X \leq F^{-1}(y)) && \text{(by (iii))}\\ &= F(F^{-1}(y))\\ &= y && \text{(by (i))}\end{aligned}$$

For $y = 0$, $P(F(X) \leq 0) = \lim_n P(F(X) \leq \frac{1}{n}) = \lim_n \frac{1}{n} = 0$.
For $y = 1$, $P(F(X) \leq 1) = \lim_n P(F(X) \leq 1 - \frac{1}{n}) = \lim_n 1 - \frac{1}{n} = 1$.
Thus, $F(X) = Y \sim Unif(0,1)$.

Exercise 1.2.5. Omit.

Exercise 1.2.6. Omit.

Exercise 1.2.7. Omit.

1.3. **Random Variables.**

Exercise 1.3.1. We want to show that $\sigma(X) = \sigma(X^{-1}(\mathcal{A}))$.

($\supset$) By definition of $\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$, it is clear.
($\subset$) It suffices to show that X is $\sigma(X^{-1}(\mathcal{A}))$-measurable. Firstly, for any σ-field $\mathcal{D}$ on Ω, $\mathcal{T} = \{B \in \mathcal{S} : \{X \in B\} \in \mathcal{D}\}$ is also a σ-field on S. Construct $\mathcal{T}$ with $\mathcal{D} = \sigma(X^{-1}(\mathcal{A}))$. Then, $\mathcal{T}$ is a σ-field containing $\mathcal{A}$. Since $\mathcal{A}$ generates $\mathcal{S}$, $\mathcal{S} = \mathcal{T}$, which means that X is $\sigma(X^{-1}(\mathcal{A}))$-measurable.

Exercise 1.3.2.

$$\begin{aligned}\{X_1 + X_2 < x\} &= \{X_1 < x - X_2\}\\ &= \bigcup_{q \in \mathbb{Q}} \{X_1 < 1 < x - X_2\}\\ &= \bigcup_{q \in \mathbb{Q}} [\{X_1 < q\} \cap \{X_2 < x - q\}] \in \mathcal{F}.\end{aligned}$$

Exercise 1.3.3. Since f is continuous, $f(X_n)$ and $f(X)$ are r.v. Moreover, on $\Omega_0 = \{\omega : X_n(\omega) \to X(\omega)\}$, $f(X_n)(\omega) \to f(X)(\omega)$, which means that $\Omega_0 \subset \Omega_0' = \{\omega : f(X_n)(\omega) \to f(X)(\omega)\}$, since f is continuous. Thus, $P(\Omega') \geq P(\Omega) = 1$, that is, $f(X_n) \to f(X)$ a.s.

Exercise 1.3.4. (i) Let f be a continuous function from $\mathbb{R}^d \to \mathbb{R}$ and let $\mathcal{O}$, $\mathcal{O}_d$ be collections of open sets in $\mathbb{R}$, $\mathbb{R}^d$, respectively. Since pre-image of open set in $\mathbb{R}$ under f is also open(by continuity), $\sigma(f^{-1}(\mathcal{O})) \subset \sigma(\mathcal{O}_d) = \mathcal{B}(\mathbb{R}^d)$. Also, by Exercise 1.3.1, $\sigma(f^{-1}(\mathcal{O})) = f^{-1}(\mathcal{B}(\mathbb{R}))$. Thus, f is $\big(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathbb{R})\big)$-measurable.

(ii) Let $\mathcal{A}$ be the smallest σ-field that makes all the continuous functions measurable. By (i), we showed that $\mathcal{A} \subset \mathcal{B}(\mathbb{R}^d)$. To show the converse, define $f_i((x_1, \ldots, x_d)) = x_i$, for $i = 1, \cdots, d$, who are continuous. Then, $f_i^{-1}((a_i, b_i)) \in \mathcal{B}(\mathbb{R}^d)$, and thus,

$$f_1^{-1}((a_1, b_1)) \cap \cdots \cap f_d^{-1}((a_d, b_d)) = (a_1, b_1) \times \cdots \times (a_d, b_d),$$

for any $a_i, b_i \in \mathbb{R}$. This means that any open boxes in $\mathbb{R}^d$ is in $\mathcal{A}$. Therefore, $\mathcal{B}(\mathbb{R}^d) \subset \mathcal{A}$.

Exercise 1.3.5. Let f be a lower semicontinuous function. f is l.s.c if and only if for any $\epsilon > 0$, there exists $\delta > 0$ such that $f(y) > f(x) - \epsilon$ for $y \in B(x, \delta)$. We want to show that f is l.s.c if and only if $\{x : f(x) \leq a\}$ is closed for each $a \in \mathbb{R}$.

($\Rightarrow$) Suppose not. Then, there exists $a_0 \in \mathbb{R}$ such that $\{x : f(x) \leq a\}$ is not closed, that is, there exists a limit point x_0 of $\{x : f(x) \leq a\}$ with $f(x_0) > a_0$. For $\epsilon = \frac{f(x_0) - a_0}{2}$, there exists $\delta > 0$ such that $f(y) > f(x_0) - \epsilon = \frac{f(x_0) + a_0}{2} > a_0$ for $y \in B(x_0, \delta)$. This contradicts that x_0 is a limit point of $\{x : f(x) \leq a\}$.

($\Leftarrow$) Fix x_0. Let's show that f is l.s.c. at x_0. For any $\epsilon > 0$, let $a = f(x_0) - \epsilon$. Then, $x_0 \in \{x : f(x) > a\}$. Since $\{x : f(x) > a\} = \{x : f(x) \leq a\}^c$ is an open set, there exists $\delta > 0$ such that $B(x_0, \delta) \subset \{x : f(x) > a\}$. That is, $f(y) > f(x) - \epsilon$ for $y \in B(x_0, \delta)$. Thus, f is l.s.c. at x_0. Since x_0 is arbitrary, f is l.s.c.

For measurability, by Exercise 1.3.1, it suffices to show that $f^{-1}((a, \infty))$ is in $\mathcal{B}(\mathbb{R})$. If f is l.s.c, then $f^{-1}((-\infty, a]) = \{x : f(x) \leq a\}$ is closed. Thus, $f \in \mathcal{B}(\mathbb{R})$.

Exercise 1.3.6. Let's show that $f^\delta = \sup\{f(y) : |y - x| < \delta\}$ is l.s.c. By Exercise 1.3.5, it suffices to show that $A_a \equiv \{x : f^\delta(x) \leq a\}$ is closed for any $a \in \mathbb{R}$. Let x_0 be a limit point of A_a. For any $y \in B(x_0, \delta)$, there exists $z_y \in A_a$ different from x_0 such that $|x_0 - z_y| < \delta - |x_0 - y|$. Since $|z_y - y| \leq |z_y - x_0| + |x_0 - y| < \delta$ and $z \in A_a$, $f(y) \leq a$. Since this holds for all $y \in B(x_0, \delta)$, $\sup\{f(y) : |y - x_0| < \delta\} \leq a$. That is, $x_0 \in A_a$, which implies that A_a is closed. Similar argument gives us that f_δ is u.s.c.

Since semicontinuous functions are measurable (Exercise 1.3.5), f^0 and f_0, limits of semicontinuous functions, are also measurable. Thus,

$$\{f^0 \neq f_0\} = \bigcup_{n=0}^{\infty} \left\{x : f^0(x) - f_0(x) > \frac{1}{n}\right\}$$

is measurable. (Clearly, $f^0 \geq f_0$)

Exercise 1.3.7. (1) Obviously, simple functions are $\mathcal{F}$-measurable. And, for any $f_n \in \mathcal{F}$, $\limsup f_n$ is also $\mathcal{F}$-measurable.

(2) Any $\mathcal{F}$-measurable function is a pointwise limit of simple functions.

By (1) and (2), class of $\mathcal{F}$-measurable functions is the smallest class containing the simple functions and closed under pointwise limit.

Exercise 1.3.8.

$$(\Omega, \sigma(X)) \xrightarrow{X} (\mathbb{R}, \mathcal{B}(\mathbb{R})) \xrightarrow{f} (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

($\Leftarrow$) Since X is $\sigma(X)$-measurable, $Y = f(X)$ is $\sigma(X)$-measurable. (Composition of measurable functions is measurable.)

($\Rightarrow$) If one can show that $\mathcal{T} = \{Y = f(X) : f \in \mathcal{B}(\mathbb{R})\}$ contains simple functions with respect to $(\Omega, \sigma(X))$ and is closed under pointwise limit, then by previous Exercise 1.3.7, $\{Y : Y \in \sigma(X)\} \subset \mathcal{T}$.

(a) Let φ be a simple function with respect to $(\Omega, \sigma(X))$. That is,

$$\varphi(\omega) = \sum_{m=1}^{n} c_m I_{A_m}(\omega) \quad \text{where disjoint } A_m \in \sigma(X).$$

Since $A_m \in \sigma(X)$, there exist disjoint $B_m \in \mathcal{B}(\mathbb{R})_X$ such that $A_m = X^{-1}(B_m)$. Thus,

$$\begin{aligned}\varphi(\omega) &= \sum_{m=1}^{n} c_m I_{A_m}(\omega)\\ &= \sum_{m=1}^{n} c_m I_{X^{-1}(B_m)}(\omega)\\ &= \sum_{m=1}^{n} c_m I_{B_m}(X(\omega))\end{aligned}$$

Therefore, $\mathcal{T}$ contains simple functions with respect to $(\Omega, \sigma(X))$.

(b) Now, let Y be a pointwise limit of $Y_n = f_n(X)$ where $f \in \mathcal{B}(\mathbb{R})$. That is, $Y(\omega) = \lim_n f_n(X(\omega))$ for all $\omega \in \Omega$. Define $f(x) \stackrel{(\star)}{=} \limsup_n f_n(x)$. Since $f_n \in \mathcal{B}(\mathbb{R})$, $f \in \mathcal{B}(\mathbb{R})$. Also, $Y(\omega) = f(X(\omega))$ for all $\omega \in \Omega$. Thus, $\mathcal{T}$ is closed under pointwise limit.

($\star$) : From the definition of f_n and f, we can ensure that f_n converges to f only on $x \in X(\Omega)$. Since we don't know whether convergence also holds on $\mathbb{R} \backslash X(\Omega)$ or not, we define $f(x)$ by limsup of $f_n(x)$.

Exercise 1.3.9. Omit.

1.4. **Integration.**

Exercise 1.4.1. Suppose not. Then, there exists n such that $\mu(A_n) > 0$, where $A_n = \{f(x) > \frac{1}{n}\}$. Since $f \geq 0$,

$$\int f d\mu \geq \int_{A_n} f d\mu \ \geq \ \frac{1}{n}\mu(A_n) > 0,$$

which is a contradiction.

Exercise 1.4.2. Define $f_n(x) = \sum_{m=1}^{\infty} \frac{m}{2^n} I_{E_{n,m}}(x)$. By MCT, $\int f_n(x) = \sum_{m=1}^{\infty} \frac{m}{2^n}\mu(E_{n,m})$. Note that if $x \in E_{n,m}$, then $x \in E_{n+1,2m} \cup E_{n+1,2m+1}$. Thus, $f_n(x) \leq f_{n+1}(x)$. And, since $f(x) - f_n(x) < \frac{1}{2^n}$, $f_n(x) \to f(x)$ for all x. Thus, by MCT, $\int f_n d\mu = \int f d\mu$. Combining two results, we get the desired result.

Exercise 1.4.3. (i) Let $g \geq 0$. Since $\int g dm = \sup\{\int s dm : 0 \leq s \leq f, s \text{ simple}\}$, there exists simple $0 \leq s \leq g$ such that $\int g dm - \epsilon < \int s dm$. Thus, $\int |g - s| dm = \int (g - s) dm < \epsilon$.

For integrable $g = g^+ - g^-$, find simple s^+ and s^- such that

$$\int |g^+ - s^+| dm < \frac{\epsilon}{2} \ \int |g^- - s^-| dm < \frac{\epsilon}{2}.$$

Then,

$$\begin{aligned} \int |g - (s^+ - s^-)| dm \leq \int |g - s^+| dm + \int |g - s^-| dm \\ < \epsilon. \end{aligned}$$

$s^+ - s^-$ is the desired simple function.

(ii) Firstly, let $\varphi = I_A$, where A is measurable. By outer regularity of Lebesgue measure, there exists open $G \supset A$ such that $m(G - A) > \frac{\epsilon}{2}$. Since $G = \biguplus_{i=1}^{\infty}(a_i, b_i)$, by continuity of measure, there exists n such that $m(G \backslash \biguplus_{i=1}^{n}(a_i, b_i)) < \frac{\epsilon}{2}$. By relabeling (a_i, b_i) for $i = 1, \ldots, n$, we can construct a step function q such that

$$\begin{aligned} \int |\varphi - q| dm &\leq \int |I_A - I_G| dm + \int |I_G - q| dm \\ &= m(G - A) + m(G \backslash \uplus_{i=1}^{n} (a_i, b_i)) \\ &< \epsilon. \end{aligned}$$

Now, let $\varphi = \sum_{k=1}^{n} b_k I_{A_k}$. For each k, there exists step function q_k such that

$$\int |I_{A_k} - q_k| dm < \frac{\epsilon}{\sum_{k=1}^{n} |b_k|}.$$

Then,

$$\begin{aligned} \int |\varphi - \sum_{k=1}^{n} b_k q_k| dm &= \int \sum_{k=1}^{n} |b_k||I_{A_k} - q_k| dm \\ &\leq \sum_{k=1}^{n} |b_k| \int |I_{A_k} - q_k| dm \end{aligned}$$

$$< \epsilon.$$

And we can construct a step function from $\sum_k b_k q_k$.

(iii) For given ϵ, pick sufficiently small positive $\delta < \frac{\epsilon}{\sum_{j=1}^k |c_j|}$. Define a continuous function r as follows.

On $x \in [a_{j-1} + \delta, a_j - \delta]$, $r(x) = c_j$.

On $x \in [a_j - \delta, a_j]$, connect $(a_j - \delta, c_j)$ and $(a_j, 0)$ continuously.

On $x \in [a_j, a_j + \delta]$, connect $(a_j, 0)$ and $(a_j + \delta, c_{j+1})$ continuously.

Then,

$$\int |q - r| d\mu < \delta \sum_{j=1}^n |c_j| < \epsilon.$$

Exercise 1.4.4. Suppose g is a step function. Then,

$$\begin{aligned}
\int g(x) \cos nx dx &= \sum_{j=1}^k c_j \int_{a_{j-1}}^{a_j} \cos nx dx \\
&= \sum_{j=1}^k c_j \frac{1}{n} (\sin(na_j) - \sin(na_{j-1})) \\
&\leq \sum_{j=1}^k \frac{2|c_j|}{n} \to 0.
\end{aligned}$$

For integrable g, previous exercise says there exists a step function q such that $\int |g - q| dx < \epsilon$, for given $\epsilon > 0$. Then,

$$\begin{aligned}
\int |g(x) \cos nx| dx &= \int |g(x) \cos nx - q(x) \cos nx + q(x) \cos nx| dx \\
&\leq \int |g - q| |\cos nx| dx + \int |q(x) \cos nx| dx \\
&\leq \int |g - q| dx + \int |q(x) \cos nx| dx \\
&< \epsilon + \int |q(x) \cos nx| dx \to \epsilon.
\end{aligned}$$

Since q is a step function, it goes to 0 as n increases. Letting $\epsilon \to 0$, we get the desired result.

1.5. **Properties of Integral.**

Exercise 1.5.1. Our claim is that if $||g||_\infty = M$, then $|g| \leq M$ *a.e.* Suppose not. Then,

$$\mu(\{x : |g(x)| > M\}) = \mu(\cup_{n=1}^\infty \{x : |g(x)| > M + \frac{1}{n}\}) > 0.$$

Thus, there exists $N \in \mathbb{N}$ such that $\mu(\{x : |g(x)| > M + \frac{1}{N}\}) > 0$, which contradicts on definition of $||g||_\infty$. Therefore, $|fg| \leq |f| ||g||_\infty$ *a.e.* By integrating, we get the desired result.

Exercise 1.5.2. As shown in the previous exercise, we know that $|f| \leq ||f||_\infty$ *a.e.* Then, $\int |f|^p d\mu \leq ||f||_\infty^p$, and thus $\limsup_p ||f||_p \leq ||f||_\infty$. For the converse, let $B_\epsilon = \{x : |f(x)| >$

$||f||_\infty - \epsilon\}$. Then, for sufficiently small $\epsilon > 0$, $0 < \mu(B_\epsilon \leq 1$. Note that

$$\begin{aligned}\int |f(x)|^p d\mu &\geq \int_{B_\epsilon} |f(x)|^p d\mu \\ &> \int_{B_\epsilon} (||f||_\infty - \epsilon)^p d\mu \\ &= \mu(B_\epsilon(||f||_\infty - \epsilon)^p\end{aligned}$$

Letting $p \to \infty$ and letting $\epsilon \to 0^+$ give us $\liminf_p ||f||_p \geq ||f||_\infty$

Exercise 1.5.3. (i)

$$||f+g||_p^p = \int |f+g|^p d\mu \leq \int 2^p(|f|^p + |g|^p) d\mu = 2^p||f||_p^p + 2^p||g||_p^p < \infty.$$

Thus, $||f+g||_p < \infty$. If $||f+g||_p = 0$, then clearly $||f+g||_p \leq ||f||_p + ||g||_p$. Suppose not. Applying Holder Inequality to $|f||f+g|^{p-1}$ and $|g||f+g|^{p-1}$ with p, $q = \frac{p}{p-1}$,

$$\begin{aligned}\int |f||f+g|^{p-1} d\mu &\leq ||f||_p \left(\int (|f+g|^{p-1})^{\frac{p}{p-1}} d\mu\right)^{\frac{p-1}{p}} \\ &= ||f||_p ||f+g||_p^{p-1} \\ \int |g||f+g|^{p-1} d\mu &\leq ||g||_p \left(\int (|f+g|^{p-1})^{\frac{p}{p-1}} d\mu\right)^{\frac{p-1}{p}} \\ &= ||g||_p ||f+g||_p^{p-1}\end{aligned}$$

Then,

$$\begin{aligned}||f+g||_p^p &= \int |f+g|^p d\mu \\ &\leq \int (|f| + |g|)|f+g|^{p-1} d\mu \\ &\leq (||f||_p + ||g||_p)||f+g||_p^{p-1}.\end{aligned}$$

Since $||f+g||_p < \infty$, we get $||f+g||_p \leq ||f||_p + ||g||_p$.

(ii) $(p=1)$ $|f+g| \leq |f| + |g|$

$(p=\infty)$ Let $||f||_\infty = M_1$ and $||g||_\infty = M_2$. Since

$$\begin{aligned}\{x : |f(x)| + |g(x)| > M_1 + M_2\} &\subset \{x : |f(x)| + |g(x)| > M_1 + M_2\} \\ &\subset \{x : |f(x)| > M_1\} \cup \{x : |g(x)| > M_2\},\end{aligned}$$

$\mu(\{x : |f(x)| + |g(x)| > M_1 + M_2\}) \leq 0$. Thus, $||f+g||_\infty \leq M_1 + M_2$.

Exercise 1.5.4. Let $f_n := \sum_{m=0}^n f\mathbf{1}_{E_m} = f\mathbf{1}_{\uplus_{m=0}^n E_m}$. Then, $f_n \to f\mathbf{1}_E$ and $|f_n| \leq f$: integrable. By DCT, we get the desired result.

Exercise 1.5.5. Let $f_n := g_n + g_1^-$. Since $g_n \uparrow g$, $-g_n^- \geq -g_1^-$. Thus, $f_n = g_n^+ - g_n^- + g_1^- \geq 0$. By MCT,

$$\int f_n d\mu = \int g_n d\mu + \int g_1^- d\mu \to \int g d\mu + \int g_1^- d\mu.$$

Since $\int g_1^- d\mu < \infty$, by substracting, we get the desired result.

Exercise 1.5.6. MCT

Exercise 1.5.7. (i) MCT

(ii) $\boxed{\textit{Step 1}}$ $g \geq 0$.

By (i), for given ϵ, there exists N such that $\int g d\mu < \int (g \wedge N) d\mu + \frac{\epsilon}{2}$. Pick $\delta = \frac{\epsilon}{2N}$. Then, for A with $\mu(A) < \delta$,

$$\int_A g d\mu < \int_A (g \wedge N) d\mu + \frac{\epsilon}{2} \leq N\mu(A) + \frac{\epsilon}{2} < \epsilon.$$

$\boxed{\textit{Step 2}}$ g : integrable.

$|g| = g^+ + g^-$. For given ϵ, there exists N such that

$$\int g^+ d\mu < \int (g^+ \wedge N) d\mu + \frac{\epsilon}{4}$$
$$\int g^- d\mu < \int (g^- \wedge N) d\mu + \frac{\epsilon}{4}.$$

Pick $\delta = \frac{\epsilon}{4N}$. Then, for A with $\mu(A) < \delta$,

$$\begin{aligned}
\int_A |g| d\mu &= \int_A g^+ d\mu + \int_A g^- d\mu \\
&< \int_A (g^+ \wedge N) d\mu + \int_A (g^- \wedge N) d\mu + \frac{\epsilon}{2} \\
&\leq 2N\mu(A) + \frac{\epsilon}{2} < \epsilon.
\end{aligned}$$

Exercise 1.5.8. Let $c \in (a, b)$. Since f is integrable, by previous exercise(1.5.7), for given $\epsilon > 0$, there exists $\delta > 0$ such that if $\mu(A) < \delta$, then $\int_A |f| d\mu < \epsilon$. For $|x - c| < \delta$,

$$|g(x) - g(c)| = \left| \int_c^x f(y) dy \right| \leq \int_c^x |f(y)| dy < \epsilon.$$

Since c is arbitrary in (a, b), g is continuous on (a, b).

Exercise 1.5.9. (i) $(1 \leq p < \infty)$

$\boxed{\textit{Step 1}}$ $f \geq 0$.

There exists $(\varphi_n)_{n \in \mathbb{N}}$ such that $0 \leq \varphi_n \leq f$ and $\varphi_n \uparrow f$. Let $g_n = |f - \varphi_n|^p$. $g_n \to 0$ and $|g_n| \leq (|f| + |\varphi_n|)^p \leq 2^p |f|^p$: integrable. By DCT, $\int g_n d\mu = \int |f - \varphi_n|^p \to 0$. Thus, $||\varphi_n - f||_p \to 0$.

$\boxed{\textit{Step 2}}$ Integrable $f = f^+ - f^-$.

There exist $(\varphi_n^+)_{n \in \mathbb{N}}$, $(\varphi_n^-)_{n \in \mathbb{N}}$ such that

$$0 \leq \varphi_n^+ \leq f^+ \text{ and } \varphi_n^+ \uparrow f^+$$
$$0 \leq \varphi_n^- \leq f^- \text{ and } \varphi_n^- \uparrow f^-.$$

For $\varphi_n \equiv \varphi_n^+ - \varphi_n^-$,

$$||\varphi_n - f||_p \leq ||\varphi_n^+ - f^+||_p + ||\varphi_n^- - f^-||_p \to 0.$$

Exercise 1.5.10. By MCT, $\int \sum_{n=1}^{\infty} |f_n| d\mu = \sum_{n=1}^{\infty} \int |f_n| d\mu < \infty$. For $g_m \equiv \sum_{n=1}^{m} f_n$, $g_m \to g \equiv \sum_{n=1}^{\infty} f_n$ and $|g_m| \le \sum_{n=1}^{m} |f_n| \le \sum_{n=1}^{\infty} |f_n|$: integrable. By DCT,

$$\int g_m d\mu = \int \sum_{n=1}^{m} f_n d\mu = \sum_{n=1}^{m} \int f_n d\mu \to \int g d\mu = \int \sum_{n=1}^{\infty} f_n d\mu.$$

1.6. **Expected Value.**

Exercise 1.6.1. For a support line $l(x) = a(x - x_0) + \varphi(x_0)$ at x_0 of $\varphi(x)$ $(l(x) \le \varphi(x))$, strict convexity implies that $(*)$ $l(x) < \varphi(x)$ if $x \neq x_0$. For $x_0 = EX$, $E\varphi(X) - El(X) = E\varphi(X) - l(EX) = E\varphi(X) - \varphi(EX) = 0$ implies $\varphi(X) = l(X)$ a.s. since $\varphi(X) \ge l(X)$. Thus, $X = EX$ a.s.
$(*)$: Suppose $l(y) = \varphi(y)$ for $y \neq x_0$. Then, with $\lambda = \frac{1}{2}$,

$$\begin{aligned} \varphi\left(\frac{x_0}{2} + \frac{y}{2}\right) &> \frac{1}{2}\varphi(x_0) + \frac{1}{2}\varphi(y) \\ &= \frac{1}{2}l(x_0) + \frac{1}{2}l(y) \\ &= l\left(\frac{x_0}{2} + \frac{y}{2}\right) \end{aligned}$$

which contradicts that $l(x)$ is a support line of $\varphi(x)$. The last equality comes from the linearity of $l(x)$.

Exercise 1.6.2. Let $l(x_1, \dots, x_n) = \varphi(EX_1, \dots, EX_n) + g_x^\top (x_1 - EX_1, \dots, x_n - EX_n)$, where $g_x \in \partial f(x)$. Then, $\varphi(x) \ge l(x)$ for any $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Taking expectation gives us the desired result.

$$E\varphi(X) \ge El(X) = \varphi(EX).$$

Exercise 1.6.3. (i) Let $\varphi(x) = x^2$ and $A = (-\infty, -a] \cup [a, \infty)$. Then, it is clear that $i_A = \inf\{\varphi(y) : y \in A\} = a^2$. By Chebyshev's inequality, we have

$$a^2 P(|X| \ge a) \le EX^2 = b^2 \Rightarrow P(|X| \ge a) \le \frac{b^2}{a^2}.$$

If some random variable X attains equality above, then X has to satisfy following identities.

$$\int_A dP = \frac{b^2}{a^2}, \int_{A^c} dP = 1 - \frac{b^2}{a^2}, \int x^2 dP = b^2.$$

It is clear that X with $P(X = a) = \frac{b^2}{a^2}$, $P(X = 0) = 1 - \frac{b^2}{a^2}$ satisfies identities. Therefore, inequality is sharp.

(ii) For $a > 0$, it's clear that if $x \in \{x : |x| \ge a\}$, then $x^2 \ge a^2$. Thus,

$$a^2 \mathbf{1}_{(|X| \ge a)} \le X^2 \mathbf{1}_{(|X| \ge a)} = X^2 - X^2 \mathbf{1}_{(|X| < a)}.$$

Since EX^2 is finite, both functions on each sides are integrable. Thus, linearity of integration and order-preserving property of integration give us

$$\int a^2 \mathbf{1}_{(|X| \ge a)} dP \le \int X^2 dP - \int X^2 \mathbf{1}_{(|X| < a)} dP.$$

We know that LHS $= a^2P(|X| \geq a)$. Let's focus on RHS. The first term in RHS is clearly EX^2. For the second term, we can apply MCT. (It is clear that conditions of MCT hold.) By MCT,

$$\begin{aligned}
\lim_{a\to\infty} a^2P(|X| \geq a) &\leq EX^2 - \lim_{a\to\infty} \int X^2\mathbf{1}_{(|X|<a)}dP \\
&= EX^2 - \int \lim_{a\to\infty} X^2\mathbf{1}_{(|X|<a)}dP \\
&= EX^2 - EX^2 = 0
\end{aligned}$$

The last equality holds since $EX^2 < \infty$. Therefore, we have

$$\lim_{a\to\infty} a^2P(|X| \geq a)/EX^2 = 0.$$

However, Chebyshev's inequality with $\varphi(x) = x^2$ and $A = (-\infty, -a] \cup [a, \infty)$, so that $i_A = a^2$, gives us following inequality.

$$a^2P(|X| \geq a)/EX^2 \leq 1,$$

which is not sharp.

Exercise 1.6.4. (i) Let $A = [a, \infty)$, $\varphi(y) = (y + b)^2$ where $0 < b < a$. It is clear that $i_A = \inf\{\varphi(y) : y \in A\} = (a + b)^2$. Then by Chebyshev's inequality,

$$(a + b)^2P(Y \geq a) \leq E[\varphi(Y)] = E[Y^2 + 2bY + b^2] = E[X^2 + 2bX + b^2] = E[\varphi(X)].$$

The second equality holds since $EY = EX, Var(Y) = Var(X)$. Note that

$$E[(X + b)^2] = p(a + b)^2.$$

Thus,

$$P(Y \geq a) \leq p.$$

If $Y = X$, then equality holds.

(ii) Let $p = \frac{\sigma^2}{(a^2+\sigma^2)}$, $b = \frac{p}{1-p}a$, $A = [a, \infty)$, and $\varphi(y) = (y + b)^2$. Note that $0 < p < 1$ and $b > 0$. Then, it is clear that $i_A = (a + b)^2$. By Chebyshev's inequality,

$$(a + b)^2P(Y \geq a) \leq E[\varphi(Y)] = E[Y^2 + 2bY + b^2] = \sigma^2 + b^2,$$

since $E(Y) = 0, Var(Y) = E(Y^2) - E(Y)^2 = E(Y^2)\sigma^2$. Thus, we have

$$P(Y \geq a) \leq \frac{\sigma^2 + b^2}{(a + b)^2}.$$

Since $b = \frac{p}{1-p}a = \frac{\sigma^2}{a}$, we have

$$P(Y \geq a) \leq \frac{\sigma^2 + b^2}{(a + b)^2} = \frac{\sigma^2 + \frac{\sigma^4}{a^2}}{(a + \frac{\sigma^2}{a})^2} = \frac{\frac{\sigma^2}{a^2}(a^2 + \sigma^2)}{\frac{1}{a^2}(a^2 + \sigma^2)^2} = \frac{\sigma^2}{(a^2 + \sigma^2)}.$$

And equality holds when Y has $P(Y = a) = p$ and $P(Y = -b) = 1 - p$, which satisfies

$$E(Y) = ap - b(1 - p) = 0$$

$$Var(Y) = E[(Y-0)^2] = a^2p + b^2(1-p) = \frac{a^2\sigma^2}{a^2+\sigma^2} + \frac{\sigma^4}{a^2+\sigma^2} = \sigma^2.$$

Exercise 1.6.5. (i) It suffices to show that $^\forall n$, $^\exists X$ such that $EX = 0$, $Var(X) = 1$ and $P(|X| > \epsilon) \le \frac{1}{n}$. Construct X as follows.

$$P(X = x) = \begin{cases} 1 - \frac{1}{n}, & \text{if } x = 0 \\ \frac{1}{2n}, & \text{if } x = \pm\sqrt{n} \end{cases}$$

Then, $EX = 0$, $Var(X) = 1$ and $P(|X| > \epsilon) \le P(X \ne 0) = \frac{1}{n}$.

(ii) Similarly,

$$P(X = x) = \begin{cases} 1 - \frac{1}{n}, & \text{if } x = 1 \\ \frac{1}{2n}, & \text{if } x = 1 \pm \sqrt{n}\sigma \end{cases}$$

Then, $EX = 1$, $Var(X) = \sigma^2$ and $P(|X| > y) \le P(X \ne 1) = \frac{1}{n}$.

Exercise 1.6.6. Let $f = Y$ and $g = I_{(Y>0)}$. Then,

$$\begin{aligned} E|fg| &= E|YI_{(Y>0)} = EYI_{(Y>0)} = EY \\ E|f|^2 &= EY^2 \\ E|g|^2 &= EI_{(Y>0)} = P(Y > 0). \end{aligned}$$

By Cauchy-Schwarz Ineq., $(E|fg|)^2 \le E|f|^2E|g|^2$. Thus, $(EY)^2 \le EY^2P(Y > 0)$.

Exercise 1.6.7. Checking (i), (ii) in Theorem 1.6.8 is obvious. For (iii),

$$\begin{aligned} Eg(X_n) &= \int |X_n|^{\frac{2}{\alpha}} dm \\ &= \int n^2 I_{(\frac{1}{n+1}, \frac{1}{n})} dm \\ &= \frac{n}{n+1} < 1 \text{ for all } n. \end{aligned}$$

Thus, Theorem 1.6.8 can be applied.
Suppose that $Y \ge |X_n| = X_n$. Then,

$$\begin{aligned} \int Y dm &\ge \sum_{n=1}^{\infty} n^\alpha \frac{1}{n(n+1)} \\ &= \sum_{n=1}^{\infty} \frac{n^{\alpha-1}}{n+1} = \infty \end{aligned}$$

The last equality holds since $0 < \alpha - 1 < 1$. Therefore, X_n are not dominated by an integrable function.

Exercise 1.6.8. $\boxed{\textit{Step 1}}$ g : Indicator function.

Let $g(x) = I_A(x)$ where $A \in \mathcal{B}(\mathbb{R})$. Then,

$$\int g(x)\mu(dx) = \int I_A(x)\mu(dx)$$

$$
\begin{aligned}
&= \mu(A) \\
\int g(x)f(x)dx &= \int I_A(x)f(x)d(x) \\
&= \int_A f(x)dx \\
&= \mu(A).
\end{aligned}
$$

$\boxed{\textit{Step 2}}$ g : simple function.
Linearity of Integral.

$\boxed{\textit{Step 3}}$ $g \geq 0$.
There exists non-negative simple function $s_n \uparrow g$. Then,

$$
\begin{aligned}
\int g(x)\mu(dx) &= \int \lim_{n\to\infty} s_n(x)\mu(dx) \\
&\overset{\text{(MCT)}}{=} \lim_{n\to\infty} \int s_n(x)\mu(dx) \\
&\overset{\textit{Step } 2}{=} \lim_{n\to\infty} \int s_n(x)f(x)dx \\
&\overset{\text{(MCT)}}{=} \int g(x)f(x)dx.
\end{aligned}
$$

$\boxed{\textit{Step 4}}$ g : Integrable.
$g = g^+ - g^-$. By *Step 3*, $\int g(x)\mu(dx) = \int g(x)f(x)dx$.

Exercise 1.6.9. Induction.

Exercise 1.6.10. Induction.

Exercise 1.6.11. (i) Let $A = (|X| > 1)$. Then, $|X|^j < |X|^k$ for $0 < j < k$ on A.

$$
\begin{aligned}
\int |X|^j dP &= \int_A |X|^j dP + \int_{A^c} |X|^j dP \\
&\leq \int_A |X|^k dP + 1 \\
&\leq \int |X|^k dP + 1 < \infty.
\end{aligned}
$$

(ii) Let $\varphi(x) = x^{\frac{k}{j}} I_{(x\geq 0)}$. It is easy to check that φ is convex. Let $Y = |X|^j$. Then, $|EY| = E|X|^j < \infty$(by (i)) and $|E\varphi(Y)| = E|X|^k < \infty$. By Jensen's Ineq.,

$$
\left(E|X^j|\right)^{\frac{k}{j}} = \varphi(EY) \leq E\varphi(Y) = E|X|^k.
$$

Thus, we get $E|X|^j \leq (E|X|^k)^{\frac{j}{k}}$.

Exercise 1.6.12. $P\left(e^X = y_m\right) = P(X = \log y_m) = p(m)$.

$$
Ee^X = \sum_{m=1}^{n} p(m)y_m
$$

$$
\begin{aligned}
e^{EX} &= \exp\left(\sum_{m=1}^{n} p(m)\log y_m\right) \\
&= \exp\left(\log \Pi_{m=1}^{n} y_m^{p(m)}\right) \\
&= \Pi_{m=1}^{n} y_m^{p(m)}.
\end{aligned}
$$

By Jensen's Ineq., $Ee^X \geq \exp(EX)$. And, we get the desired result.

Exercise 1.6.13. Exercise 1.5.5

Exercise 1.6.14. (i)

$$E(1/X; X > y) = E(1/X; 1/X < 1/y) \leq \frac{1}{y}P(1/X < 1/y) = \frac{1}{y}P(X > y).$$

Thus, $\lim_{y\to\infty} yE(1/X; X > y) \leq \lim_{y\to\infty} P(X > y) = 0$.

(ii) Let $0 < y < \epsilon$.

$$
\begin{aligned}
yE(1/X; X > y) &= E(y/X; X > y \text{ and } 0 \leq X < \epsilon) + E(y/X; X > y \text{ and } \epsilon \leq X). \\
&\leq P(X > y, 0 \leq X < \epsilon) + E(y/X; \epsilon \leq X). \\
&\leq P(0 < X < \epsilon) + E(y/X; \epsilon \leq X).
\end{aligned}
$$

For the second term on right hand side, $\frac{y}{X}I_{\epsilon\leq X} \leq \frac{y}{\epsilon}I_{\epsilon\leq X} \leq 1$ and $\frac{y}{X}I_{\epsilon\leq X} \to 0$ as $y \to 0$. Thus, by DCT,

$$\limsup_{y\to 0} yE(1/X; X > y) \leq P(0 < X < \epsilon).$$

Letting $\epsilon \to 0$, we get the desired result.

Exercise 1.6.15. Exercise 1.5.6

Exercise 1.6.16. Let $X_k = XI_{\biguplus_{n=0}^k A_n}$. Then, $X_k \to XI_A$ and $|X_k| \leq |X|$: integrable. By DCT, $EX_k \to E(X; A)$. Note that

$$EX_k = \int_{\biguplus_{n=0}^k A_n} XdP = \sum_{n=0}^{k}\int_{A_n} XdP = \sum_{n=0}^{k} E(X; A_n).$$

Thus, $E(X; A) = \lim_{k\to\infty} EX_k = \sum_{n=0}^{\infty} EX_k$.

1.7. **Product Measures, Fubini's Theorem.**

Exercise 1.7.1. Since $|f(x,y)| \geq 0$, by Fubini's Theorem, $\int_{X\times Y} |f|d\mu = \int_X \int_Y |f(x,y)|\mu_2(dy)\mu_1(dx) < \infty$. And by Fubini's Theorem again, we get the desired result.

Exercise 1.7.2. Construct a probability space $(Y, \mathcal{B}, \nu) = ([0,\infty), \mathcal{B}([0,\infty)), \lambda)$, where λ denotes a Lebesgue measure. Let $f(x,y) = I_{\{(x,y)\,:\,0\leq y<g(x)\}}$. Then,

$$
\begin{aligned}
\int_X \int_Y f d\lambda d\mu &= \int_X g d\mu \\
\int_{X\times Y} f d(\mu\times\lambda) &= (\mu\times\lambda)(\{(x,y) \,:\, 0 \leq y < g(x)\})
\end{aligned}
$$

$$\int_Y \int_X f d\mu d\lambda = \int_0^\infty \mu(\{x : g(x) > y\} dy.$$

Since $f \geq 0$, Fubini's Theorem implies that

$$\int_X \int_Y f d\lambda d\mu = \int_{X\times Y} f d(\mu \times \lambda) = \int_Y \int_X f d\mu d\lambda.$$

And we get the desired result.

Exercise 1.7.3. (i) Fubini's Theorem with $f(x, y) = I_{\{(x,y):a<x\leq y\leq b\}} \geq 0$.

(ii)

$$\begin{aligned}
&\int_{(a,b]} F(y)dG(y) + \int_{(a,b]} G(x)dF(x) \\
&= \int_{(a,b]} \{F(y) - F(a)\}dG(y) + \int_{(a,b]} \{G(x) - G(a)\}dF(x) + \int_{(a,b]} F(a)dG(y) + \int_{(a,b]} G(a)dF(x) \\
&\overset{\text{(i)}}{=} \int I_{\{(x,y):a<x\leq y\leq b\}} d(\mu \times \nu) + \int I_{\{(x,y):a<y\leq x\leq b\}} d(\mu \times \nu) + F(a)(G(b) - G(a)) + G(a)(F(b) - F(a)) \\
&= (\mu \times \nu)((a, b] \times (a, b]) + (\mu \times \nu)(\{(x, y) : a < x = y \leq b\} + F(a)G(b) + F(b)G(a) - 2F(a)G(a) \\
&= (F(b) - F(a))(G(b) - G(a)) + \sum_{x\in(a,b]} \mu(\{x\})\nu(\{x\}) + F(a)G(b) + F(b)G(a) - 2F(a)G(a) \\
&= F(b)(G(b) - F(a)G(a) + \sum_{x\in(a,b]} \mu(\{x\})\nu(\{x\}).
\end{aligned}$$

(iii) If F=G, then the third term in right hand side of (ii) equals zero. Thus,

$$\int_{(a,b]} 2F(y)dF(y) = F^2(b) - F^2(a).$$

Exercise 1.7.4. Let $f = I_{\{(x,y):x<y\leq x+c\}}$. Then,

$$\begin{aligned}
\int_{\mathbb{R}} \int_{\mathbb{R}} f d\mu(y)dx &= \int_{\mathbb{R}} (F(x + c) - F(x))dx \\
\int_{\mathbb{R}} \int_{\mathbb{R}} f dx d\mu(y) &= \int_{\mathbb{R}} c d\mu(y) \\
&= c\mu(\mathbb{R}).
\end{aligned}$$

Since $f \geq 0$, Fubini's Theorem gives us that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f d\mu(y)dx = \int_{\mathbb{R}} \int_{\mathbb{R}} f dx d\mu$$

and we get the desired result.

Exercise 1.7.5.

$$\begin{aligned}
\int_0^a \int_0^\infty |e^{-xy} \sin x| dy dx &= \int_0^a \int_0^\infty e^{-xy} |\sin x| dy dx \\
&= \int_0^a \frac{|\sin x|}{x} dx \\
&\leq \int_0^a 1 dx < \infty.
\end{aligned}$$

By Exercise 1.7.1, $e^{-xy}\sin x I_{\{(x,y):0<x<a,0<y\}}$ is integrable and we can apply Fubini's Theorem.

$$\begin{aligned}\int_0^a \frac{\sin x}{x}dx = \int_0^a\int_0^\infty e^{-xy}\sin x dydx &= \int_0^\infty\int_0^a e^{-xy}\sin x dxdy\\ &= \int_0^\infty \left[\frac{e^{-xy}(-y\sin x-\cos x)}{1+y^2}\right]_0^a dy\\ &= \int_0^\infty \frac{1-e^{-ay}(y\sin a+\cos a)}{1+y^2}dy\\ &= \frac{\pi}{2}-\cos a\int_0^\infty\frac{e^{-ay}}{1+y^2}dy-\sin a\int_0^\infty\frac{ye^{-ay}}{1+y^2}dy.\end{aligned}$$

The last equality comes from the fact that $\int\frac{1}{1+y^2}dy=\tan^{-1}(y)$. Moreover,

$$\begin{aligned}\left|\int_0^a\frac{\sin x}{x}dx-\frac{\pi}{2}\right| &= \left|\cos a\int_0^\infty\frac{e^{-ay}}{1+y^2}dy+\sin a\int_0^\infty\frac{ye^{-ay}}{1+y^2}dy\right|\\ &\le \int_0^\infty e^{-ay}\frac{|\cos a|+y|\sin a|}{1+y^2}dy\\ &\le \int_0^\infty e^{-ay}\frac{1+y}{1+y^2}dy\\ &\le \int_0^\infty e^{-ay}\cdot 2dy=\frac{2}{a}.\end{aligned}$$

Letting $a\to\infty$, we get

$$\int_0^\infty\frac{\sin x}{x}dx=\frac{\pi}{2}.$$

Note that since $\frac{\sin x}{x}$ is not integrable on $\mathbb{R}_{\ge 0}$, we can't calculate its integration directly.

2. Laws of Large Numbers

2.1. Independence.

Exercise 2.1.1. By Thm 2.1.8, it suffices to show that $P(X_1\le z_1,\dots,X_n\le z_n)=\Pi_{i=1}^n P(X_i\le z_i)$. Let $c_i=\int_{-\infty}^\infty g_i(x_i)dx_i$, for $i=1,\dots,n$. Then, we have $c_1\cdots c_n=1$. Our claim is that

$$P(X_i\le z_i)=\frac{1}{c_i}\int_{-\infty}^{z_i}g_i(x_i)dx_i.$$

WLOG, let $i=1$.

$$\begin{aligned}P(X_1\le z_1) &= P(-\infty<X_1\le z_1,-\infty<X_2<\infty,\dots,-\infty<X_n<\infty)\\ &= (c_2\cdots c_n)\int_{-\infty}^{z_1}g_1(x_1)dx_1\\ &= \frac{1}{c_1}\int_{-\infty}^{z_1}g_1(x_1)dx_1.\end{aligned}$$

Now,

$$P(X_1\le z_1,\dots,X_n\le z_n)=\left(\int_{-\infty}^{z_1}g_1(x_1)dx_1\right)\cdots\left(\int_{-\infty}^{z_n}g_n(x_n)dx_n\right)$$

$$= \left(\frac{1}{c_1}\int_{-\infty}^{z_1} g_1(x_1)dx_1\right)\cdots\left(\frac{1}{c_n}\int_{-\infty}^{z_n} g_n(x_n)dx_n\right)$$
$$= \prod_{i=1}^{n} P(X_i \le z_i)$$

Exercise 2.1.2. Let $\mathcal{A}_i = \{X_i^{-1}(x_i) : x_i \in S_i\} \cup \{\emptyset, \Omega\}$. Let $I = \{n_1, \ldots, n_k\} \subset \{1, \ldots, n\}$ and $J = \{1, \ldots, n\} \setminus I = \{m_1, \ldots, m_{n-k}\}$. Then,

$$P(X_{n_1} = x_{n_1}, \cdots, X_{n_k} = x_{n_k}) = \sum_{x_{m_1} \in S_{m_1}} \cdots \sum_{x_{m_{n-1}} \in S_{m_{n-k}}} P(X_1 = x_1, \cdots, X_n = x_n)$$
$$= \prod_{i \in I} P(X_i = x_i).$$

By Lemma 2.1.5, $\mathcal{A}_i$'s are independent. By Theorem 2.1.7, $\sigma(\mathcal{A}_i)$'s are independent. Sine $X_i^{-1}(A) \in \sigma(\mathcal{A}_i)$ for all $A \in 2^{S_i}$, $\sigma(X_i) \subset \sigma(\mathcal{A}_i)$. Therefore, $X_1, \ldots, X_n$ are independent.

Exercise 2.1.3. (1) $h(\rho(x,y)) \to [0,\infty)$

(i) $h(\rho(x,y)) = 0 \iff \rho(x,y) = 0 \iff x = y$.
First $\iff$ holds since $h(0) = 0$ and $h'(x) > 0$ for $x > 0$.

(ii) $h(\rho(x,y)) = h(\rho(y,x))$ by symmetry of metric ρ.

(iii)
$$h(\rho(x,z)) = \int_0^{\rho(x,z)} h'(u)du$$
$$\le \int_0^{\rho(x,y)+\rho(y,z)} h'(u)du$$
$$\le \int_0^{\rho(x,y)} h'(u)du + \int_0^{\rho(y,z)} h'(u)du$$
$$= h(\rho(x,y)) + h(\rho(y,z)).$$

The first inequality holds since ρ is metric. The second inequality holds since h' is decreasing.

(2) By simple calculus, one can show that $h(x) = \frac{x}{x+1}$ satisfies (1).

Exercise 2.1.4. It is easy to check the uncorrelatedness. To show that X_n's are not independent, it suffices to show that there exist A and B such that $P(X_n \in A, X_m \in B) \ne P(X_n \in A)P(X_m \in B)$ for different m and n.

For $0 < \epsilon < 1$, let $\begin{cases} A = X_n^{-1}\big((0,\epsilon)\big) \\ B = X_m^{-1}\big(X_m(A)\big) \end{cases}$. Then, $A \subset B$. Note that we have $2n$ intervals in A. Let l_ϵ be the length of one interval. Now, pick ϵ so that $2nl\epsilon < \frac{1}{m}$. This is possible since $P(A) \searrow 0$ as $\epsilon \searrow 0$. Then, it is clear that $P(B) < 1$. Thus, $P(A \cap B) = P(A) > P(A)P(B)$. Therefore, X_n and X_m are not independent.

Exercise 2.1.5. (i) Let D_μ be a set of discontinuities of μ. By Exercise 1.2.3, D_μ is at most countable.

$$P(X + Y = 0) = \int\int \mathbf{1}_{(x+y=0)} d\mu d\nu$$

$$= \int \mu(\{-y\})d\nu$$
$$= \int_{D_\mu} \mu(\{-y\})d\nu + \int_{D_\mu^c} \mu(\{-y\})d\nu.$$

Note that for $-y \in D_\mu^c$, $\mu(\{-y\}) = P(X \le -y) - P(X < -y) = 0$. Thus,

$$P(X+Y=0) = \int_{\mathbb{R}} \mu(\{-y\})\mathbf{1}_{D_\mu}(y)d\nu$$
$$= \sum_{y \in D_\mu} \mu(\{-y\})\nu(\{y\}).$$

The second equality holds by Exercise 1.5.6.

(ii) With $Y = -Y$ in (i), $P(X = Y) = 0$ since $D_\mu = \emptyset$.

Exercise 2.1.6. Note that

$$\sigma(f(X)) = \{X^{-1}(f^{-1}(B)) : B \in \mathcal{B}(\mathbb{R})\}$$
$$\sigma(f(Y)) = \{Y^{-1}(g^{-1}(B)) : B \in \mathcal{B}(\mathbb{R})\}$$

Since f and g are measurable, $f^{-1}(B)$ and $g^{-1}(B)$ are in $\mathcal{B}(\mathbb{R})$ for any $B \in \mathcal{B}(\mathbb{R})$. Thus, $\sigma(f(X)) \subset \sigma(X)$ and $\sigma(g(Y)) \subset \sigma(Y)$. Therefore, independence of X and Y implies the independence of $f(X)$ and $g(Y)$.

Exercise 2.1.7. By Exercise 2.1.2, it suffices to show that $P(Z_i = \alpha, Z_j = \beta) = P(Z_i = \alpha)P(Z_j = \beta)$ for all $\alpha, \beta \in \{0, 1, \ldots, K-1\}$.
Firstly, for given $Y = y$, $P(Z_i = \alpha | Y = y) = \frac{1}{K}$ since K is a prime. Thus,

$$P(Z_i = \alpha) = \sum_{y=0}^{K-1} P(Z_i = \alpha | Y = y)P(Y = y)$$
$$= K \cdot \frac{1}{K^2} = \frac{1}{K}.$$

Now, for given $Y = y$, $Z_i = \alpha$ and $Z_j = \beta$ means that $X + iy \equiv \alpha$ and $X + jY \equiv \beta$. That is,

$$(i-j)y \equiv \alpha - \beta \text{ (mod } K).$$

Since $gcd(i-j, K) = 1$, there exists unique $y_0 \in \{0, 1, \ldots, K-1\}$ satisfying $Z_i = \alpha$ and $Z_j = \beta$. And, with y_0, $x_0 \in \{0, 1, \ldots, K-1\}$ uniquely determined. Thus,

$$P(Z_i = \alpha, Z_j = \beta) = \frac{1}{K^2}.$$

Therefore,

$$P(Z_i = \alpha, Z_j = \beta) = \frac{1}{K^2} = \frac{1}{K} \cdot \frac{1}{K} = P(Z_i = \alpha)P(Z_j = \beta) \text{ for all } \alpha, \beta \in \{0, 1, \ldots, K-1\}.$$

Exercise 2.1.8. $X_1, X_2, X_3 \overset{i.i.d}{\sim} \frac{1}{2}\mathbf{1}_{\{1\}} + \frac{1}{2}\mathbf{1}_{\{-1\}}$ and $X_4 = X_1X_2X_3$.

Exercise 2.1.9. $\mathcal{A}_1 = \Big\{\{1,2\}, \Omega\Big\}$ and $\mathcal{A}_2 = \Big\{\{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \Omega\Big\}$

Exercise 2.1.10.

$$P(X+Y=n)=\sum_m P(X=m,Y=n-m)=\sum_m P(X=m)P(Y=n-m)$$

The last equality holds by the independence.

Exercise 2.1.11. Omit.

Exercise 2.1.12. Omit.

Exercise 2.1.13. (a) $F(z)=\frac{1}{9}\mathbf{1}_{[0,1)}(z)+\frac{2}{9}\mathbf{1}_{[1,2)}(z)+\frac{1}{3}\mathbf{1}_{[2,3)}(z)+\frac{2}{9}\mathbf{1}_{[3,4)}(z)+\frac{1}{9}\mathbf{1}_{[4,\infty)}(z)$
(b) For some $c\in\left[0,\frac{2}{9}\right]$

	0	1	2
0	$\frac{1}{9}$	c	$\frac{2}{9}-c$
1	$\frac{2}{9}-c$	$\frac{1}{9}$	c
2	c	$\frac{2}{9}-c$	$\frac{1}{9}$

Exercise 2.1.14.

$$\begin{aligned}P(XY\le z)&=\int\int\mathbf{1}_{(xy\le z)}dFdG\\&=\int\int\left(\mathbf{1}_{(x\le z/y,y\neq 0)}+\mathbf{1}_{(0\le z,y=0)}\right)dFdG\\&=\int_{\mathbb{R}\backslash\{0\}}F\left(\frac{z}{y}\right)dG+P(Y=0)I(z\ge 0).\end{aligned}$$

Exercise 2.1.15. For simplicity, let $\Omega=[0,1)$. We have

$$Y_n(\omega)=\begin{cases}0, \text{ if } \omega\in\left\{\left[\frac{2m}{2^n},\frac{2m+1}{2^n}\right):m=0,1,\dots,2^{n-1}-1\right\}\\ 1, \text{ if } \omega\in\left\{\left[\frac{2m+1}{2^n},\frac{2m+2}{2^n}\right):m=0,1,\dots,2^{n-1}-1\right\}\end{cases}.$$

By construction, it is obvious that $P(Y_n=0)=P(Y_n=1)=\frac{1}{2}$. To show that $Y_1,Y_2,\dots$ are independent, we have to show that $Y_1,\dots,Y_n$ are independent for all $n\in\mathbb{N}$. By Exercise 2.1.2, it suffices to show that $P(Y_1=y_1,\dots,Y_n=y_n)=\prod_{i=1}^n P(Y_i=y_i)$, where $y_i\in\{0,1\}$. By induction, one can show that $\{\omega:Y_1(\omega)=y_1,\dots,Y_n(\omega)=y_n\}=\left[\frac{l}{2^n},\frac{l+1}{2^n}\right)$ for some $l\in\{0,1,\dots,2^n-1\}$ for any y_i's. Therefore, $P(Y_1=y_1,\dots,Y_n=y_n)=\frac{1}{2^n}=\prod_{i=1}^n P(Y_i=y_i)$, where $y_i\in\{0,1\}$.

2.2. **Weak Laws of Large Numbers.**

Exercise 2.2.1. We assume that $EX_i^2<\infty$ so that $\mathrm{Var}(X_i)<\infty$ for all i. Since L^2 convergence implies convergence in probability, it suffices to show that

$$E|S_n/n-\nu_n|^2\to 0 \text{ as } n\to\infty.$$

Note that

$$E|S_n/n-\nu_n|^2=\frac{1}{n^2}E|S_n-ES_n|^2$$

$$
\begin{aligned}
&= \frac{1}{n^2}\,\mathrm{Var}(S_n) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) \\
&\le \frac{1}{n}\sum_{i=1}^{n}\frac{\mathrm{Var}(X_i)}{i}
\end{aligned}
$$

The last equality holds since X_i's are uncorrelated. Now, fix $\epsilon > 0$. Since $\frac{\mathrm{Var}(X_n)}{n} \to 0$, there exists $N \in \mathbb{N}$ such that $\frac{\mathrm{Var}(X_n)}{n} < \epsilon$ for all $n > N$. Let $M = \sum_{i=1}^{N}\frac{\mathrm{Var}(X_i)}{i} < \infty$. Then, for $n > N$,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\frac{\mathrm{Var}(X_i)}{i} &< \frac{1}{n}\left(M + \sum_{i=N+1}^{n}\epsilon\right) \\
&\le \frac{1}{n}\cdot M + \epsilon
\end{aligned}
$$

By taking limsup and letting $\epsilon \to 0$, we get

$$E|S_n/n - \nu_n|^2 \to 0,$$

which is the desired result.

Exercise 2.2.2. We assume that $r(0)$ is finite. Let $S_n = X_1 + \cdots + X_n$. Since $ES_n = 0$,

$$
\begin{aligned}
E\left[\left(\frac{S_n}{n}\right)^2\right] &= \mathrm{Var}\left(\frac{S_n}{n}\right) \\
&= \frac{1}{n^2}\sum_{1\le i,j\le n} EX_iX_j \\
&\le \frac{1}{n^2}\Big(\sum_{i=1}^{n}EX_i^2 + 2\sum_{i=1}^{n-1}EX_iX_{i+1} + 2\sum_{i=1}^{n-2}EX_iX_{i+2} + \cdots + 2\sum_{i=1}^{2}EX_iX_{i+n-2} + 2EX_1X_n\Big)
\end{aligned}
$$

Now, since $r(k) \to 0$ as $k \to \infty$, for given $\epsilon > 0$, $^\exists N$ such that for $n \ge N$, $|r(n)| < \epsilon$. Moreover, since $EX_i^2 \le r(0) < \infty$, by Cauchy-Schwartz Inequality, $E|X_iX_j| \le (EX_i^2)^{1/2}(EX_j^2)^{1/2} \le r(0) < \infty$. Then, for $n \ge N+1$, we have

$$
\begin{aligned}
&E\left[\left(\frac{S_n}{n}\right)^2\right] \\
&\quad\le \frac{1}{n^2}\Big(\sum_{i=1}^{n}EX_i^2 + 2\sum_{i=1}^{n-1}E|X_iX_{i+1}| + \cdots + 2\sum_{i=1}^{n-N+1}E|X_iX_{i+N-1}| + 2\sum_{i=1}^{n-N}EX_iX_{i+N} + \cdots + 2EX_1X_n\Big) \\
&\quad\le \frac{1}{n^2}\Big(nr(0) + 2(n-1)r(0) + \cdots + 2(n-N+1)r(0) + 2\sum_{i=1}^{n-N}r(N) + 2\sum_{i=1}^{n-N-1}r(N+1) + \cdots + 2r(n-1)\Big) \\
&\quad\le \frac{1}{n^2}\Big(2Nnr(0) + 2\sum_{i=1}^{n-N}|r(N)| + 2\sum_{i=1}^{n-N-1}|r(N+1)| + \cdots + 2|r(n-1)|\Big)
\end{aligned}
$$

$$\leq \frac{1}{n^2}\Big(2Nnr(0) + 2\sum_{i=1}^{n-N} i \cdot \epsilon\Big)$$
$$\leq \frac{1}{n^2}\Big(2Nnr(0) + 2\sum_{i=1}^{n} i \cdot \epsilon\Big)$$
$$\leq \frac{2Nr(0)}{n} + \frac{n(n+1)}{n^2}\epsilon$$

Since $\epsilon > 0$ is arbitrary, $E\left[\left(\frac{S_n}{n}\right)^2\right] \to 0$, which means that $S_n/n \to 0$ in L^2. Therefore, $S_n/n \to 0$ in probability, which is the desired result.

Exercise 2.2.3. (i) Since U_i's are *i.i.d.* and f is a measurable function, $f(U_i)$'s are also *i.i.d.* Note that since U_i follows Uniform distribution on [0,1], $Ef(U_1) = \int_0^1 f(x)dx$. By assumption, $E|f(U_1)| = \int_0^1 |f(x)|dx < \infty$. By WLLN,

$$I_n = \frac{f(U_1) + \cdots + f(U_n)}{n} \to Ef(U_1) = \int_0^1 f(x)dx = I \text{ in prob.}$$

(ii) Since $\int_0^1 |f(x)|^2 dx < \infty$,

$$\mathrm{Var}(f(U_1)) = Ef(U_1)^2 - (Ef(U_1))^2 = \int_0^1 f(x)^2 dx - \left(\int_0^1 f(x)dx\right)^2 < \infty.$$

By Chebyshev's inequality with $\varphi(x) = x^2$ and $A = \{x \in \mathbb{R} : |x| > a/n^{1/2}\}$, we have $i_A = \inf\{\varphi(y) : y \in A\} = a^2/n$ and

$$\begin{aligned} P(|I_n - I| > a/n^{1/2}) &\leq \frac{n}{a^2} E(I_n - I)^2 \\ &= \frac{n}{a^2}\,\mathrm{Var}(I_n) \\ &= \frac{1}{a^2}\,\mathrm{Var}(f(U_1)) \\ &= \frac{1}{a^2}\left[\int_0^1 f(x)^2 dx - \left(\int_0^1 f(x)dx\right)^2\right] \end{aligned}$$

Exercise 2.2.4. (1) $E|X_i| = \infty$.

Since $\sum_{k=2}^{\infty} \frac{1}{k\log k} \geq \int_2^\infty \frac{1}{x\log x}dx = \big[\log(\log x)\big]_2^\infty = \infty$, $E|X_i| = \sum_{k=2}^\infty \frac{C}{k\log k} = \infty$.

(2) There exists $\mu < \infty$ so that $S_n/n \to \mu$ in probability.

Since

$$nP(|X_1| > n) \leq n \sum_{k=n+1}^{\infty} \frac{C}{k^2 \log k} \leq \frac{n}{\log n} \sum_{k=n+1}^{\infty} \frac{C}{k^2} \leq \frac{n}{\log n}\int_n^\infty \frac{C}{x^2}dx = \frac{C}{\log n},$$

$nP(|X_1| > n) \to 0$ as $n \to \infty$. Also, there exists $\mu < \infty$ such that

$$\mu_n := E(X_1 \mathbf{1}_{(|X_1| \leq n)}) = \sum_{k=2}^{\infty} (-1)^k \cdot k \cdot \frac{C}{k^2\log k} = \sum_{k=2}^{\infty}(-1)^k \frac{C}{k\log k} \to \mu$$

since it is alternating series. Therefore, by Theorem 2.2.12, $S_n/n \to \mu$ in probability.

Exercise 2.2.5. By definition, X_i takes only positive values.

(1) There exists a sequence $\mu_n \to \infty$ such that $S_n/n - \mu_n \to 0$ in probability.
Firstly, for $x \geq 2$, $xP(|X_i| > x) = xP(X_i > x) = \frac{2}{\log x} \to 0$ as $x \to \infty$. Let $\mu_n = E(X_1 \mathbf{1}_{(|X_1| \leq n)})$. Then, for $n \geq 3$,

$$\mu_n = \int_e^n \frac{e}{x \log x} dx = [e \log(\log x)]_e^n = e \log(\log n) \to \infty.$$

Therefore, by Theorem 2.2.12, $S_n/n - \mu_n \to 0$ in probability.

(2) $E|X_i| = \infty$.
Since $E|X_i| = EX_i \geq \mu_n$ for all n, $E|X_i| = \infty$.

Exercise 2.2.6. (i) $X = \sum_{n=1}^{\infty} \mathbf{1}_{(X \geq n)}$.

(ii)
$$\begin{aligned}
EX^2 &= \sum_{x=1}^{\infty} x^2 P(X = x) \\
&= P(X = 1) + (1+3)P(X = 2) + \cdots + (1 + 3 + \cdots + (2n-1))P(X = n) + \cdots \\
&= P(X \geq 1) + 3P(X \geq 2) + \cdots + (2n-1)P(X \geq n) + \cdots \\
&= \sum_{n=1}^{\infty} (2n-1)P(X \geq n).
\end{aligned}$$

Exercise 2.2.7.

$$\begin{aligned}
\int_{-\infty}^{\infty} h(y)P(X \geq y)dy &= \int_{-\infty}^{\infty} \int_{\Omega} h(y) \mathbf{1}_{(X \geq y)}(\omega) dP dy \\
&= \int_{\Omega} \int_{-\infty}^{\infty} h(y) \mathbf{1}_{(X \geq y)}(\omega) dP dy \\
&= \int_{\Omega} H(X) dP = EH(X).
\end{aligned}$$

The second equality holds by Fubini.

Exercise 2.2.8. Let's use Theorem 2.2.11 in Textbook with $X_{n,k} = X_k$ for $k = 1, \ldots, n$ and $b_n = 2^{m(n)}$ where $m(n) = \min\left\{m : 2^{-m} m^{-3/2} \geq n^{-1}\right\}$. Let's check the conditions.

(0) $b_n > 0$ with $b_n \to \infty$.
By definition of $m(n)$.

(i) $\sum_{k=1}^{n} P(|X_{n,k}| > b_k) \to 0$.

$$\begin{aligned}
\sum_{k=1}^{n} P(|X_{n,k}| > b_k) = nP(|X_1| > 2^{m(n)}) &= n \sum_{j=m(n)+1}^{\infty} p_j \\
&= \sum_{j=m(n)+1}^{\infty} \frac{n}{2^j \cdot j(j+1)} \leq \sum_{j=m(n)+1}^{\infty} \frac{n}{2^j \cdot m(n)^2} \\
&= \frac{n}{2^{m(n)} m(n)^2} \leq m(n)^{-\frac{1}{2}} \to 0
\end{aligned}$$

(ii) $b_n^{-2}\sum_{k=1}^n E\bar{X}_{n,k}^2 \to 0$.

$$\begin{aligned}
b_n^{-2}\sum_{k=1}^n E\bar{X}_{n,k}^2 &= \frac{n}{2^{2m(n)}} E\left(X_1^2 \mathbf{1}_{\{|X_1|\le 2^{m(n)}\}}\right) \\
&\le \frac{m(n)^{3/2}}{2^{m(n)}} E\left(X_1^2 \mathbf{1}_{\{|X_1|\le 2^{m(n)}\}}\right) \\
&= \frac{m(n)^{3/2}}{2^{m(n)}} \left(\sum_{k=1}^{m(n)} \left(2^k-1\right)^2 p_k + p_0\right). \qquad (1)
\end{aligned}$$

Note that

$$\begin{aligned}
\left(\sum_{k=1}^{m(n)} \left(2^k-1\right)^2 p_k + p_0\right) &\le 1 + \sum_{k=1}^{m(n)} 2^{2k} p_k \\
&= 1 + \sum_{k=1}^{m(n)} 2^k \cdot \frac{1}{k(k+1)} \\
&= 1 + \sum_{1\le k<\frac{m(n)}{2}} 2^k \cdot \frac{1}{k(k+1)} + \sum_{\frac{m(n)}{2}\le k\le m(n)} 2^k \cdot \frac{1}{k(k+1)} \\
&\le 1 + \sum_{1\le k<\frac{m(n)}{2}} 2^k + \frac{4}{m(n)\cdot(m(n)+2)} \sum_{\frac{m(n)}{2}\le k\le m(n)} 2^k \cdot \frac{1}{k(k+1)} \\
&= 1 + 2^{\frac{m(n)}{2}+1} - 2 + \frac{4}{m(n)\cdot(m(n)+2)} \left(2^{m(n)+1} - 2^{\frac{m(n)}{2}+1}\right) \\
&\le \frac{2^{m(n)}}{m(n)^2} \cdot C \quad \text{for large } n.
\end{aligned}$$

Thus,

$$(1) \le \frac{m(n)^{\frac{3}{2}}}{2^{m(n)}} \cdot \frac{2^{m(n)}}{m(n)^2} \cdot C = C \cdot m(n)^{-\frac{1}{2}} \to 0$$

By (0)-(ii), Theorem 2.2.11 gives us that $\frac{S_n - a_n}{b_n} \xrightarrow{p} 0$, where $a_n = \sum_{k=1}^n E\bar{X}_{n,k}$.

Now,

$$\begin{aligned}
a_n &= nE(X_1 \mathbf{1}_{\{|X_1|\le 2^{m(n)}\}} \\
&= n\left[\sum_{k=1}^{m(n)} (2^k-1)p_k - p_0\right] \\
&= n\left[-\frac{1}{m(n)+1} + \sum_{k=m(n)+1}^{\infty} \frac{1}{2^k \cdot k(k+1)}\right] \\
&\le n\left[-\frac{1}{m(n)+1} + \frac{1}{m(n)^2} \sum_{k=m(n)+1}^{\infty} \frac{1}{2^k}\right]
\end{aligned}$$

$$\sim -\frac{n}{m(n)}.$$

Since $n \le 2^{m(n)} m(n)^{3/2} \le Cn$, $\log n \sim m(n)$. Thus,

$$a_n \sim \frac{n}{\log n}$$
$$b_n = 2^{m(n)} \sim n(m(n))^{-3/2} \sim n(\log n)^{-3/2}.$$

Thus,

$$\frac{S_n - a_n}{b_n} \sim \frac{S_n + \frac{n}{\log n}}{n(\log n)^{-\frac{3}{2}}} \xrightarrow{p} 0.$$

Since $(\log n)^{-\frac{1}{2}} \to 0$,

$$\frac{S_n + n(\log n)^{-1}}{n(\log n)^{-1}} \xrightarrow{p} 0.$$

Finally,

$$\frac{S_n}{n/\log n} \xrightarrow{p} 0.$$

Exercise 2.2.9.

2.3. **Borel-Cantelli Lemmas.**

Exercise 2.3.1. Since $\cup_{k\ge n} A_k \supset A_m$ for all $m \ge n$, $P(\cup_{k\ge n} A_k) \ge \sup_{k\ge n} P(A_k)$. Thus,

$$P(\limsup A_n) = P(\cap_{n=1}^{\infty} \cup_{k\ge n} A_k) = \lim_{n\to\infty} P(\cup_{k\ge n} A_k) \ge \lim_{n\to\infty} \sup_{k\ge n} P(A_k) = \limsup P(A_n).$$

By applying above result to $B_n = A_n^c$, we get

$$P(\liminf A_n) \le \liminf P(A_n).$$

Exercise 2.3.2. $\epsilon > 0$ given. Since f is uniformly continuous on $[-M-1, M+1]$, there exists $\delta \in (0,1)$ such that if $x, y \in [-M-1, M+1]$ and $|x-y| < \delta$, then $|f(x) - f(y)| < \epsilon$. Then, $|f(x) - f(y)| \ge \epsilon$ implies $|x - y| \ge \delta$ or $|x| \ge M$. Thus,

$$P(|f(X_n) - f(X)| \ge \epsilon) \le P(|X_n - X| \ge \delta) + P(|X| \ge M).$$

Since $X_n \to X$ in probability,

$$\limsup_n P(|f(X_n) - f(X)| \ge \epsilon) \le P(|X| \ge M).$$

By letting $M \to \infty$, we get

$$\lim_n P(|f(X_n) - f(X)| \ge \epsilon) = 0.$$

Exercise 2.3.3. We know that

$$P(l_n = k) = \begin{cases} \frac{1}{2^{k+1}}, & k = 0, 1, \dots, n-1 \\ \frac{1}{2^n}, & k = n. \end{cases}$$

(1) $\limsup \frac{l_n}{\log_2 n} = 1$ a.s.

(i) $\limsup \frac{l_n}{\log_2 n} \le 1$ a.s.
Fix $\epsilon > 0$. Let $A_n = \{l_n > (1+\epsilon)\log_2 n\}$ and $m = \lfloor (1+\epsilon)\log_2 n \rfloor$. Then,

$$\begin{aligned} P(A_n) &= 1 - P(l_n \le (1+\epsilon)\log_2 n \\ &= 1 - \sum_{k=1}^{n} \frac{1}{2^{k+1}} = \left(\frac{1}{2}\right)^{m+1} \\ &\le \left(\frac{1}{2}\right)^{(1+\epsilon)\log_2 n} = n^{-(1+\epsilon)}. \end{aligned}$$

Since $1+\epsilon > 1$, we get $\sum_{n=1}^{\infty} P(A_n) < \infty$. By Borel-Cantelli Lemma 1,

$$P(A_n \ i.o.) = P\left(\frac{l_n}{\log_2 n} > 1+\epsilon \ i.o.\right) = 0,$$

that is,

$$P\left(\frac{l_n}{\log_2 n} \le 1+\epsilon \ eventually\right) = 1.$$

Since $\epsilon > 0$ was arbitrarily chosen, we get $\limsup \frac{l_n}{\log_2 n} \le 1$ a.s.

(ii) $\limsup \frac{l_n}{\log_2 n} \ge 1$ a.s.
Construct a sequence $(r_n)_{n\in\mathbb{N}}$ by $r_1 = 1$, $r_2 = 2$, $r_3 = r_{n-1} + [\log_2 n]$ for $r \ge 3$. Let $B_n = \{X_m = 1, \text{ for all } m \text{ such that } r_{n-1} \le m \le r_n\}$. Then,

$$P(B_n) = \left(\frac{1}{2}\right)^{r_n - r_{n-1}} = \left(\frac{1}{2}\right)^{[\log_2 n]} \ge \left(\frac{1}{2}\right)^{\log_2 n} = \frac{1}{n}.$$

Since $\sum_{n=1}^{\infty} P(B_n) = \infty$ and B_n's are independent, by Borel-Cantelli Lemma 2, $P(B_n \ i.o.) = 1$. Since $B_n \subset \{l_{r_n} \ge [\log_2 n]\}$, $P(l_{r_n} \ge [\log_2 n] \ i.o.) = 1$. Note that $r_n \le n\log_2 n$ for $n \ge 2$. (It can be shown easily by induction.) Then, since $\left\{\frac{l_{r_n}}{\log_2 r_n} \ge \frac{[\log_2 n]}{\log_2 r_n}\right\} \subset \left\{\frac{l_{r_n}}{\log_2 r_n} \ge \frac{[\log_2 n]}{\log_2(n\log_2 n)} = \frac{[\log_2 n]}{\log_2 n + \log_2\log_2 n}\right\}$, we get

$$P\left(\frac{l_{r_n}}{\log_2 r_n} \ge \frac{[\log_2 n]}{\log_2 n + \log_2\log_2 n} \ i.o.\right) = 1.$$

Since $\lim_{n\to\infty} \frac{[\log_2 n]}{\log_2 n + \log_2\log_2 n} = 1$,

$$\limsup \frac{l_n}{\log_2 n} \ge 1 \ a.s.$$

(2) $\liminf \frac{l_n}{\log_2 n} = 0$ a.s.
It suffices to show that $P(l_n = 0 i.o.) = 1$. Note that $\{l_n = 0\} = \{X_n = 0\}$ are independent and $P(X_n = 0) = \frac{1}{2}$. Thus, by Borel-Cantelli Lemma 2, $P(l_n = 0 \ i.o.) = 1$.

Exercise 2.3.4. Firstly, find a subsequence $(X_{n(m)})_{m\in\mathbb{N}}$ such that $EX_{n(m)} \to \liminf EX_n$. Since $X_n \to X$ in probability, there exists a further subsequence $X_{n(m_k)}$ such that $X_{n(m_k)} \to X$ a.s. by Theorem 2.3.2 in Textbook. Then, by Fatou's Lemma,

$$EX = E(\liminf X_{n(m_k)}) \le \liminf EX_{n(m_k)} = \liminf EX_n.$$

The inequality comes from Fatou's Lemma.

Exercise 2.3.5. (a) For all subsequence X_{n_k}, there exists a sub-subsequence $X_{m(n_k)}$ converging almost surely to X. Since $|X_{m(n_k)}| \le Y$: integrable, by DCT, $EX_{m(n_k)} \to EX$. This means that every subsequence (EX_{n_k}) of (EX_n) has a sub-subsequence $EX_{m(n_k)}$ converging to EX. Therefore, $EX_n \to EX$.

(b) See Theorem 1.6.8.

Exercise 2.3.6. (a) Note that $f(x) = \frac{x}{1+x}$ is non-decreasing, $0 \le f(x) \le 1$ for all $x \ge 0$ and $f(x) = 0$ if and only if $x = 0$.

Firstly, $d(X,Y) = E(f(|X-Y|)) \ge [0,\infty)$.

(i)

$$\begin{aligned} d(X,Y) = E(f(|X-Y|)) = 0 &\iff f(|X-Y|) = 0 \text{ a.s.} \\ &\iff |X-Y| = 0 \text{ a.s.} \end{aligned}$$

(ii) Clear.

(iii) Since $|X-Z| \le |X-Y| + |Y-Z|$,

$$\begin{aligned} \frac{|X-Z|}{1+|X-Z|} &= f(|X-Z|) \\ &\le f(|X-Y| + |Y-Z|) \\ &= \frac{|X-Y| + |Y-Z|}{1+|X-Y|+|Y-Z|} \\ &= \frac{|X-Y|}{1+|X-Y|} + \frac{|Y-Z|}{1+|Y-Z|} \end{aligned}$$

Therefore,

$$d(X,Z) = E(f(|X-Z|)) \le E(f(|X-Y|)) + E(f(|Y-Z|)) = d(X,Y) + d(Y,Z).$$

(b) ($\Rightarrow$) Fix $\epsilon > 0$. Let $\varphi(x) = \frac{x}{1+x} \ge 0$ and $A = (\epsilon, \infty)$. Then, we have $i_A = \inf\{\varphi(y) : y \in A\} = \frac{\epsilon}{1+\epsilon}$. By Chebyshev's inequality, we get

$$P(|X_n - X| > \epsilon) \le \frac{1+\epsilon}{\epsilon} \cdot E\left(\frac{|X_n - X|}{1+|X_n - X|}\right) \le \frac{1+\epsilon}{\epsilon} \cdot d(X_n, X) \to 0.$$

Thus, $X_n \to X$ in probability.

($\Leftarrow$) For any $\epsilon > 0$,

$$\begin{aligned} d(X_n, X) &= E\left(\frac{|X_n - X|}{1+|X_n - X|}\right) \\ &= E\left(\frac{|X_n - X|}{1+|X_n - X|}; |X_n - X| > \epsilon\right) + E\left(\frac{|X_n - X|}{1+|X_n - X|}; |X_n - X| \le \epsilon\right) \\ &\le E(1; |X_n - X| > \epsilon) + E\left(\frac{\epsilon}{1+\epsilon}; |X_n - X| \le \epsilon\right) \\ &\le P(|X_n - X| < \epsilon) + \frac{\epsilon}{1+\epsilon}. \end{aligned}$$

Since $X_n \to X$ in probability, $\limsup_n d(X_n, X) = \frac{\epsilon}{1+\epsilon}$. Since $\epsilon > 0$ was arbitrarily chosen, by letting $\epsilon \to 0$, we get $\lim_n d(X_n, X) = 0$.

Exercise 2.3.7. By Theorem 2.3.2 in Textbook, it suffices to show that there exists a r.v. X_∞ such that for any subsequence $X_{n(m)}$, there exists a further subsequence $X_{n(m_k)}$ that converges almost surely to X_∞. Since $d(X_m, X_n) \to 0$, there exists $N_k \in \mathbb{N}$ such that $d(X_m, X_n) < 2^{-k}$ if $m, n > N_k$. Now, fix a subsequence $X_{n(m)}$. Choose increasing subsequence (m_k) so that $n(m_k) \geq N_k$. Using Chebyshev's Inequality with $\varphi(y) = \frac{y}{1+y}$ and $A = \left(\frac{1}{k^2}, \infty\right)$, we get

$$P\left(\left|X_{n(m_k)} - X_{n(m_{k-1})}\right| > \frac{1}{k^2}\right) \leq \frac{k^2+1}{2^k}.$$

Since $\sum_{k=1}^\infty = P\left(\left|X_{n(m_k)} - X_{n(m_{k-1})}\right| > \frac{1}{k^2}\right) \leq \sum_{k=1}^\infty \frac{k^2+1}{2^k} < \infty$, by Borel-Cantelli Lemma 1,

$$P\left(\left|X_{n(m_k)} - X_{n(m_{k-1})}\right| > \frac{1}{k^2} \ i.o.\right) = 0,$$

that is, $P\left(\left|X_{n(m_k)} - X_{n(m_{k-1})}\right| \leq \frac{1}{k^2} \ eventually\right) = 1$. For $\omega \in \Omega_0 := \left\{\left|X_{n(m_k)} - X_{n(m_{k-1})}\right| \leq \frac{1}{k^2} \ eventually\right\}$, $\left(X_{n(m_k)}(\omega)\right)_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Thus, there exists $X_\infty(\omega) := \lim_k X_{n(m_k)}(\omega)$. Since $P(\Omega_0) = 1$, $X_{n_{m(k)}} \to X_\infty$ a.s.

It remains to check the well-definedness of X_∞. Consider other subsequence $X_{n'(m')}$ such that $X_{n'(m'_k)} \to X'_\infty$. Since $d(X_{n(m_k)}, X_{n'(m'_k)}) \to 0$, by DCT, $d(X_\infty, X'_\infty) = 0$. By Exercise 2.3.6, it's equivalent to that $X_\infty = X'_\infty$.

Exercise 2.3.8. Since $P(\cup_n A_n) = 1$, $P(\cap_n A_n^c) = 0$. Moreover, since A_n's are independent events, so are A_n^c. And this implies that

$$\begin{aligned} P\left(\bigcap_{n=1}^\infty A_n^c\right) &= \prod_{n=1}^\infty P(A_n^c) \\ &= \prod_{n=1}^\infty (1 - P(A_n)) = 0. \end{aligned}$$

Since $P(A_n) < 1$, $\prod_{n=1}^{m-1}(1 - P(A_n)) > 0$ for $m = 2, 3, \cdots$. Thus, we have for all $m \in \mathbb{N}$,

$$P\left(\bigcap_{n=m}^\infty A_n^c\right) = 0.$$

Since $\{A_n^c \ eventually\} = \bigcup_{m=1}^\infty \bigcap_{n=m}^\infty A_n^c$,

$$\begin{aligned} P(A_n^c \ eventually) &= P\left(\bigcup_{m=1}^\infty \bigcap_{n=m}^\infty A_n^c\right) \\ &\leq \sum_{m=1}^\infty P\left(\bigcap_{n=m}^\infty A_n^c\right) = 0. \end{aligned}$$

Then, $P(A_n \ i.o.) = 1$. Thus, by Borel-Cantelli Lemma 1(Theorem 2.3.1), $\sum_{n=1}^\infty P(A_n) = \infty$.

Exercise 2.3.9. (i) Let $B_n = A_n^c \cap A_{n+1}$. Then,

$$P(\cup_{m\geq n} A_m) \leq P(A_n) + \sum_{m=n}^{\infty} P(B_m) \to 0 \quad \text{as } n \to \infty.$$

Thus, $P(A_n \text{ i.o.}) = \lim_{n\to} P(\cup_{m\geq n} A_m) = 0$.

(ii) $A_n = \left[0, \frac{1}{n}\right)$ with Lebesgue measure.

Exercise 2.3.10. 10

Exercise 2.3.11. (i) For $0 < \epsilon < 1$, $P(|X_n| > \epsilon) = p_n \to 0 \iff p_n \to 0$.

(ii) Let $A_n = \{|X_n| > \epsilon\}$. Then, $P(A_n \text{ i.o.}) = 0 \iff \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} p_n < \infty$. Note that $(\Rightarrow)$ holds since X_n's are independent.

Exercise 2.3.12. Let's enumerate $\Omega = \{\omega_1, \omega_2, \cdots\}$. Let $\Omega_0 = \{\omega_i : P(\omega_i) > 0\}$. Since $P(\Omega_0^c) = 0$, if one can show that $X_n(\omega) \to X(\omega)$ for $\omega \in \Omega_0$, then it's done. Fix $\epsilon > 0$. Since $X_n \to X$ in probability, for each $\omega_i \in \Omega_0$, there exists $N_i \in \mathbb{N}$ such that $P(|X_n - X|) > \epsilon) < \frac{1}{2}P(\omega_i)$ if $n > N_i$. Since $\frac{1}{2}P(\omega_i) < P(\omega_i)$,

$$\omega_i \notin \{\omega : |X_n - X| > \epsilon \text{ for } n > N_i\}.$$

Thus, $X_n(\omega_i) \to X(\omega_i)$ as $n \to \infty$.

Exercise 2.3.13. Pick c_n such that $P\left(|X_n| > \frac{|c_n|}{n}\right) \leq \frac{1}{2^n}$. Since $\sum_{n=1}^{\infty} P\left(\left|\frac{X_n}{c_n}\right| > \frac{1}{n}\right) < \infty$, by Borel-Cantelli Lemma 1, $P\left(\left|\frac{X_n}{c_n}\right| > \frac{1}{n} \textit{ i.o.}\right) = 0$. That is, $P\left(\left|\frac{X_n}{c_n}\right| \leq \frac{1}{n} \textit{ eventually}\right) = 1$. Now, for given $\epsilon > 0$, choose $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. Then, since $\{|X_n| \leq |c_n|\epsilon\} \supset \{|X_n| \leq |c_n|/N\}$,

$$P\left(\left|\frac{X_n}{c_n}\right| \leq \epsilon \textit{ eventually}\right) \geq P\left(\left|\frac{X_n}{c_n}\right| \leq \frac{1}{N} \textit{ eventually}\right) = 1.$$

Since ϵ is arbitrarily chosen, $X_n/c_n \to 0$ a.s.

Exercise 2.3.14. $(\Rightarrow)$ Suppose $\sum_n P(X_n > A) = \infty$ for all A. Then, by Borel-Cantelli Lemma 2, $P(X_n > A \textit{ i.o.}) = 1$ for all A. That is, $\sup X_n = \infty$ a.s. which is the contradiction.

$(\Leftarrow)$ Since $\sum_n P(X_n > A) < \infty$ for some A, by Borel-Cantelli Lemma 1, $P(X_n > A \textit{ i.o.}) = 0$. Thus, $\sup X_n \leq A < \infty$ a.s.

Exercise 2.3.15. (i) $\limsup_{n\to\infty} \frac{X_n}{\log n} = 1$ a.s.

$(\leq)$ It suffices to show that $P\left(\frac{X_n}{\log n} > 1 + \epsilon \text{ i.o.}\right) = 0$ for any $\epsilon > 0$. By definition, $P\Big(X_n > (1+\epsilon)\log n\Big) = n^{-(1+\epsilon)}$. Since $1 + \epsilon > 1$, we have

$$\sum_{n=1}^{\infty} P\Big(X_n > (1+\epsilon)\log n\Big) = \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}} < \infty.$$

By Borel-Cantelli Lemma 1, $P\Big(X_n > (1+\epsilon)\log n \textit{ i.o.}\Big) = 0$.

($\geq$) It suffices to show that $P\left(\frac{X_n}{\log n} > 1 \ i.o.\right) = 1$. Since $X_n's$ are independent and $\sum_{n=1}^{\infty} P(X_n > \log n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$, by Borel-Cantelli Lemma 2, we get

$$P(X_n > \log n \text{ i.o.}) = 1.$$

(ii) If one can show that

$$1 \overset{\text{(a)}}{\leq} \liminf_{n\to\infty} \frac{M_n}{\log n} \leq \limsup_{n\to\infty} \frac{M_n}{\log n} \overset{\text{(b)}}{\leq} 1 \quad \text{a.s.},$$

then it's done.

(a) It suffices to show that $P\left(\frac{M_n}{\log n} < 1-\epsilon \text{ i.o.}\right) = 0$ for any $\epsilon > 0$. Note that

$$\begin{aligned} P\left(\frac{M_n}{\log n} < 1-\epsilon\right) &= P\left(\max_{1\leq m\leq n} X_m < (1-\epsilon)\log n\right) \\ &= \prod_{m=1}^{n} P(X_m < (1-\epsilon)\log n) \\ &= \left(1 - \frac{1}{n^{1-\epsilon}}\right)^n \\ &\leq e^{-n^{\epsilon}}. \end{aligned}$$

The second equality holds since X_i's are *i.i.d.* By L'Hopital's rule, $e^{-n^{\epsilon}}/\frac{1}{n^2} \to 0$. Thus, by comparison test,

$$\sum_{n=1}^{\infty} P\left(\frac{M_n}{\log n} < 1-\epsilon\right) < \infty.$$

By Borel-Cantelli Lemma 1, $P\left(\frac{M_n}{\log n} < 1-\epsilon \ i.o.\right) = 0$.

(b) It suffices to show that $P\left(\frac{M_n}{\log n} < 1+\epsilon \text{ eventually}\right) = 1$ for any $\epsilon > 0$.

Fix $\epsilon > 0$. Let $\Omega_0 = \{\omega : \limsup_{n\to\infty} \frac{X_n(\omega)}{\log n} = 1\}$.

($\star$) For $\omega \in \Omega_0$, there exists $N_1(\omega) \in \mathbb{N}$ such that $\frac{X_n(\omega)}{\log n} < 1+\epsilon$ if $n > N_1(\omega)$.

Now, for $n > N_1(\omega)$,

$$\begin{aligned} \frac{M_n(\omega)}{\log n} &= \max_{1\leq m\leq n} \frac{X_n(\omega)}{\log n} \\ &= \max\left\{\max_{1\leq m\leq N_1(\omega)} \frac{X_m(\omega)}{\log n}, \max_{N_1(\omega)<m\leq n} \frac{X_n(\omega)}{\log n}\right\}. \end{aligned}$$

- There exists $N_2(\omega) \in \mathbb{N}$ such that $\max_{1\leq m\leq N_1(\omega)} \frac{X_m(\omega)}{\log n} < 1+\epsilon$ if $n > N_2(\omega)$.
- By ($\star$), $\frac{X_m(\omega)}{\log n} \leq \frac{X_m(\omega)}{\log m} < 1+\epsilon$ for $N_1(\omega) < m \leq n$.

Thus, for $n > N(\omega) = \max\{N_1(\omega), N_2(\omega)\}$, we have

$$\frac{M_n(\omega)}{\log n} < 1+\epsilon.$$

That is, $\Omega_0 \subset \left\{\frac{M_n}{\log n} < 1+\epsilon \ \textit{eventually}\right\}$. Since $P(\Omega_0) = 1$ by (i),

$$P\left(\frac{M_n}{\log n} < 1+\epsilon \ \textit{eventually}\right) = 1.$$

Exercise 2.3.16. (1) $\sum_n (1 - F(\lambda_n)) < \infty$.

Let $B_n = \{X_n > \lambda_n\}$. Since $1 - F(\lambda_n) = P(X_n > \lambda_n)$, by Borel-Cantelli Lemma 1, $P(X_n > \lambda_n \ \textit{i.o.}) = 0$. That is, $P(X_n \le \lambda_n \ \textit{eventually}) = 1$. Note that ($\star$) for $\omega \in \Omega_0 := \{\omega : X_n(\omega) \le \lambda_n \ \textit{eventually}\}$, there exists $N_1(\omega) \in \mathbb{N}$ such that $X_n(\omega) \le \lambda_n$ if $n > N_1(\omega)$.

Now, for $n > N_1(\omega)$,

$$\max_{1\le m\le n} X_m(\omega) = \max\left\{\max_{1\le m\le N_1(\omega)} X_m(\omega), \max_{N_1(\omega)<m\le n} X_m(\omega)\right\}.$$

- Since $\lambda_n \uparrow \infty$, there exists $N_2(\omega) \in \mathbb{N}$ such that $\max_{1\le m\le N_1(\omega)} X_m(\omega) \le \lambda_n$ if $n > N_2(\omega)$.
- By ($\star$), $X_m(\omega) \le \lambda_m \le \lambda_n$ for $N_1(\omega) < m \le n$.

Thus, for $n > N(\omega) = \max\{N_1(\omega), N_2(\omega)\}$, we have

$$\max_{1\le m\le n} X_m(\omega) \le \lambda_n.$$

That is, $\Omega_0 \subset \{\max_{1\le m\le n} X_m \le \lambda_n \ \textit{eventually}\}$.

Since $P(\Omega_0) = 1$, $P(\max_{1\le m\le n} X_m \le \lambda_n \ \textit{eventually}) = 1$, which means that

$$P\left(\max_{1\le m\le n} X_m > \lambda_n \ \textit{i.o.}\right) = 0.$$

(2) $\sum_n (1 - F(\lambda_n)) = \infty$.

By Borel-Cantelli Lemma 2, $P(X_n > \lambda_n \ \textit{i.o.}) = 1$. Then, clearly, $P(\max_{1\le m\le n} X_n > \lambda_n \ \textit{i.o.}) = 1$.

Exercise 2.3.17. The answers are

(i) $\mathbb{E}|Y_1| < \infty$,
(ii) $\mathbb{E}Y_1^+ < \infty$,
(iii) $n\mathbb{P}(|Y_1| > n) \to 0$, and
(iv) $\mathbb{P}(|Y_1| < \infty) = 1$.

(i) If $\mathbb{E}|Y_1| < \infty$, then

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\frac{Y_1}{n}\right| > \epsilon\right) = \sum_{n=1}^{\infty} \mathbb{P}\left(|Y_1| > \epsilon n\right) \le \int_0^{\infty} \mathbb{P}\left(|Y_1| > t\right) = \mathbb{E}|Y_1| < \infty$$

since $\mathbb{P}(|Y_1| > \cdot)$ decreases. By Borel-Cantelli, we have the convergence. Conversely, if $\mathbb{E}|Y_1| = \infty$, then

$$\infty = \mathbb{E}|Y_1| = \int_0^{\infty} \mathbb{P}(|Y_1| > t)\, dt \le \sum_{n=0}^{\infty} \mathbb{P}(|Y_1| > n)$$

since $\mathbb{P}(|Y_1| > \cdot)$ decreases. By the second Borel-Cantelli, the convergence fails.

(ii) If $\mathbb{E}Y_1^+ < \infty$, then $Y_n^+/n \to 0$ a.s. from (i). Thus for a given $\epsilon > 0$, there exists N such that $Y_n/n^+ < \epsilon$ for all $n > N$. Then,

$$\frac{\max_{m\le n} Y_m^+}{n} \le \max\left\{\frac{Y_1^+}{n}, \cdots, \frac{Y_N^+}{n}, \epsilon\right\} \to 0$$

almost surely as $n \to \infty$. That is,

$$\lim_{n\to\infty} \frac{\max_{m\le n} Y_m^+}{n} = 0$$

almost surely. Now we have

$$\limsup_{n\to\infty} \frac{\max_{m\le n} Y_m}{n} \le \limsup_{n\to\infty} \frac{\max_{m\le n} Y_m^+}{n} = 0.$$

We claim that

$$\limsup_{n\to\infty} \frac{\max_{m\le n} Y_m}{n} \ge 0.$$

Note that

$$\mathbb{P}\left(\frac{\max_{m\le n} Y_m}{n} < -\epsilon\right) = F(-n\epsilon)^n$$

where F is the distribution of Y_1. Let M be large so that $F(-n\epsilon) < \delta < 1$ for all $n > M$. Then we have

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}\left(\frac{\max_{m\le n} Y_m}{n} < -\epsilon\right) &\le \sum_{n=1}^{M} \mathbb{P}\left(\frac{\max_{m\le n} Y_m}{n} < -\epsilon\right) + \sum_{n=M+1}^{\infty} \delta^n \\
&\le M + \frac{\delta^{M+1}}{1-\delta} \\
&< \infty.
\end{aligned}$$

Therefore, by Borel-Cantelli,

$$\limsup_{n\to\infty} \frac{\max_{m\le n} Y_m}{n} \ge -\epsilon$$

almost surely. This proves our claim.

Conversely, if $\mathbb{E}Y_1^+ = \infty$, then

$$\infty = \mathbb{E}Y_1^+ = \int_0^\infty \mathbb{P}(Y_1 > t)\,dt \le \sum_{n=0}^{\infty} \mathbb{P}(Y_1 > n),$$

and thus $Y_n/n > 1$ infinitely often.

(iii) If

(iv) By definition, $Y_n/n \to 0$ in probability if and only if

$$\mathbb{P}(|Y_n| > n\epsilon) \to 0$$

for any $\epsilon > 0$. This is equivalent to $\mathbb{P}(|Y_1| = \infty) = 0$.

Exercise 2.3.18 and 2.3.19 will be solved by the way we did in Theorem 2.3.9 in Textbook.

Exercise 2.3.18. Step 1. $\frac{X_n}{an^\alpha} \to 1$ in probability.

Fix $\epsilon > 0$. Since $\beta < 2\alpha$,

$$P\left(\left|\frac{X_n - EX_n}{EX_n}\right| > \epsilon\right) \le \frac{\text{Var}(X_n)}{\epsilon^2 (EX_n)^2} \le \frac{Bn^\beta}{\epsilon^2 (EX_n)^2} \sim \frac{B}{a^2\epsilon^2} \cdot n^{\beta - 2\alpha} \to 0.$$

Thus, $\frac{X_n - EX_n}{EX_n} \to 0$ in probability. Therefore, $\frac{X_n}{EX_n} \to 1$ in probability.

Step 2. Find subsequence (n_k).

Let $n_k = k^{[2/(2\alpha-\beta)]+1}$ (+1 to avoid the case where $\left[\frac{2}{2\alpha-\beta}\right] = 0$.) and $T_k = X_{n_k}$.

Then,

$$\begin{aligned} P\left(\left|\frac{T_k - ET_k}{ET_k}\right| > \epsilon\right) &\le \frac{\text{Var}(T_k)}{\epsilon^2 (ET_k)^2} \le \frac{Bn_k^\beta}{\epsilon^2 (ET_k)^2} \sim \frac{B}{a^2\epsilon^2} \cdot n_k^{\beta-2\alpha} \\ &\le \frac{B}{a^2\epsilon^2} \cdot k^{\left(\frac{2}{2\alpha-\beta}\right)\cdot(\beta-2\alpha)} \\ &= Ck^{-2} \end{aligned}$$

The last inequality holds by construction of n_k. Since $\sum_{k=1}^\infty P\left(\left|\frac{T_k - ET_k}{ET_k}\right| > \epsilon\right) < \infty$, by Borel-Cantelli Lemma 1,

$$P\left(\left|\frac{T_k - ET_k}{ET_k}\right| > \epsilon \ i.o.\right) = 0.$$

That is, $\frac{T_k}{ET_k} \to 1$ a.s.

Step 3. $\frac{X_n}{an^\alpha} \to 1$ a.s.

For sufficiently large n such that $n_k \le n < n_{k+1}$ and $\omega \in \Omega_0 := \{\omega : \frac{T_k}{ET_k} \to 1\}$, we have

$$\frac{ET_k}{E_{T_{k+1}}} \cdot \frac{T_k(\omega)}{ET_k} = \frac{T_k(\omega)}{ET_{k+1}} \le \frac{X_n(\omega)}{EX_n} \le \frac{T_{k+1}(\omega)}{ET_k} = \frac{T_{k+1}(\omega)}{ET_{k+1}} \cdot \frac{ET_{k+1}}{ET_k}.$$

Note that

$$\frac{ET_{k+1}}{ET_k} \sim \frac{an_{k+1}^\alpha}{an_k^\alpha} = \left[\left(1 + \frac{1}{k}\right)^{\left[\frac{2}{2\alpha-\beta}\right]+1}\right]^\alpha \to 1.$$

Therefore, we have $\frac{X_n(\omega)}{EX_n} \to 1$. Since $P(\Omega_0) = 1$ by Step 2, we conclude that $\frac{X_n}{EX_n} \to 1$ a.s.

Step 4. $\frac{X_n}{n^\alpha} \to a$ a.s.

$$\frac{X_n}{n^\alpha} = a \cdot \left(\frac{EX_n}{an^\alpha}\right) \cdot \left(\frac{X_n}{EX_n}\right) \to a \quad \text{a.s.}$$

Exercise 2.3.19. Suppose $Y \sim Poisson(\lambda)$, where $\lambda > 1$. Then, we can decompose Y into $Y = Y_1 + \cdots + Y_n$, where Y_i's are independent and $Y_i \sim Poisson(\lambda_i$. Thus, without loss of generality, we may assume that X_n are independent Poisson r.v.'s with $EX_n = \lambda_n$, where $\lambda_n \le 1$.

Step 1. $\frac{S_n}{ES_n} \to 1$ in probability.

Fix $\epsilon > 0$. Since $\sum_n \lambda_n = \infty$,

$$P\left(\left|\frac{S_n - ES_n}{ES_n}\right| > \epsilon\right) \le \frac{\mathrm{Var}(S_n)}{\epsilon^2 (ES_n)^2} = \frac{\sum_n \lambda_n}{\epsilon^2 (\sum_n \lambda_n)^2} = \frac{1}{\epsilon^2 \sum_n \lambda_n} \to 0.$$

Thus, $\frac{S_n - ES_n}{ES_n} \to 0$ in probability. Therefore, $\frac{S_n}{ES_n} \to 1$ in probability.

Step 2. Find subsequence (n_k).

Since $\sum_n \lambda_n = \infty$, one can find subsequence (n_k) such that $n_k = \inf\{n : \sum_{i=1}^n \lambda_i \ge k^2\}$. Let $T_k = S_{n_k}$. Then,

$$P\left(\left|\frac{T_k - ET_k}{ET_k}\right| > \epsilon\right) \le \frac{\mathrm{Var}(T_k)}{\epsilon^2 (ET_k)^2} = \frac{1}{\epsilon^2 \sum_{i=1}^{n_k} \lambda_i} \le \frac{1}{\epsilon^2 k^2}.$$

Since $\sum_{k=1}^\infty P\left(\left|\frac{T_k - ET_k}{ET_k}\right| > \epsilon\right) < \infty$, by Borel-Cantelli Lemma 1,

$$P\left(\left|\frac{T_k - ET_k}{ET_k}\right| > \epsilon \ i.o.\right) = 0.$$

That is, $\frac{T_k}{ET_k} \to 1$ a.s.

Step 3. $\frac{S_n}{ES_n} \to 1$ a.s.

For $n \in \mathbb{N}$ such that $n_k \le n < n_{k+1}$, we have

$$k^2 \le \sum_{i=1}^{n_k} \lambda_i = ET_k \le k^2 + 1 \le (k+1)^2 \le ET_{k+1} \le (k+1)^2 + 1.$$

Note that the assumption that $\lambda_n \le 1$ is used here.

With such n and $\omega \in \Omega_0 := \{\omega : \frac{T_k(\omega)}{ET_k} \to 1\}$,

$$\frac{ET_k}{E_{T_{k+1}}} \cdot \frac{T_k(\omega)}{ET_k} = \frac{T_k(\omega)}{ET_{k+1}} \le \frac{S_n(\omega)}{ES_n} \le \frac{T_{k+1}(\omega)}{ET_k} = \frac{T_{k+1}(\omega)}{ET_{k+1}} \cdot \frac{ET_{k+1}}{ET_k}.$$

Note that

$$\frac{ET_{k+1}}{ET_k} \le \frac{(k+1)^2 + 1}{k^2} = 1 + \frac{2(k+1)}{k^2} \to 1.$$

Therefore, we have $\frac{S_n(\omega)}{ES_n} \to 1$. Since $P(\Omega_0) = 1$ by Step 2, we conclude that $\frac{S_n}{ES_n} \to 1$ a.s.

Exercise 2.3.20. It suffices to show that for any $M > 0$, $P\left(\frac{X_n}{n \log_2 n} \ge M \ i.o.\right) = 1$. Fix $M > 0$. Let $k_n = \lfloor \log_2(M \cdot n \log_2 n) \rfloor + 1$. Then, for $n \ge 2$,

$$\begin{aligned} P(X_1 \ge M \cdot n \log_2 n) &= \sum_{j=k_n}^\infty \frac{1}{2^j} = \frac{1}{2^{k_n - 1}} \\ &= \frac{1}{2^{\lfloor \log_2(M \cdot n \log_2 n) \rfloor}} \ge \frac{1}{M \cdot n \log_2 n}. \end{aligned}$$

Then, by Exercise 2.2.4,

$$\sum_{n=2}^{\infty} P\left(\frac{X_1}{n\log_2 n} \geq M\right) \geq \sum_{n=2}^{\infty} \frac{1}{M\cdot n\log_2 n} = \infty.$$

Since X_n's are independent, by Borel-Cantelli Lemma 2, we get $P\left(\frac{X_n}{n\log_2 n} \geq M \;\; i.o.\right) = 1$.

2.4. **Strong Law of Large Numbers.**

Exercise 2.4.1. $X_i, Y_i \geq 0$. $X_i \overset{i.i.d}{\sim} F$, $Y_i \overset{i.i.d}{\sim} G$. $EX_1, EY_1 < \infty$.
Let $T_n := (X_1+Y_1)+\cdots+(X_n+Y_n)$ and $N_t := \min\{n : T_n \leq t\}$. Note that $N_t \to \infty$ as $t \to \infty$. Then,

$$\frac{\sum_{i=1}^{N_t} X_i}{t} \leq \frac{R_t}{t} \leq 1 - \frac{\sum_{i=1}^{N_t} Y_i}{t}.$$

By SLLN and Theorem 2.4.7, we have

$$\frac{\sum_{i=1}^{N_t} X_i}{N_t} \overset{\text{a.s.}}{\to} EX_1, \qquad \frac{\sum_{i=1}^{N_t} Y_i}{N_t} \overset{\text{a.s.}}{\to} EY_1, \qquad \frac{N_t}{t} \overset{\text{a.s.}}{\to} \frac{1}{EX_1+EY_1}.$$

Thus, we get

$$\frac{R_t}{t} \overset{\text{a.s.}}{\to} \frac{EX_1}{EX_1+EY_1}.$$

Exercise 2.4.2. Let $Y_{n+1} = X_{n+1}/|X_n|$. Since Y_n is independent of $X_1, \cdots, X_n$, $(Y_n)_n$ is an iid sequence. Note that

$$\frac{\log|X_n|}{n} = \frac{\log|Y_1|+\cdots+\log|Y_n|}{n}$$

converges to $\mathbb{E}\log|Y_1|$ whenever it exists by the Strong Law of Large Numbers. Now to find the expectation,

$$\mathbb{E}\log|Y_1| = \int_{B_1} \frac{1}{\pi}\log|x|\,dx = \int_0^1 \int_{\partial B_r} \frac{1}{\pi}\log r\,dS\,dr = \int_0^1 2r\log r\,dr = -\frac{1}{2}.$$

Remark. For d-dimensional vectors, one can derive that it converges to $-1/d$ by a similar argument. $\square$

Exercise 2.4.3. (i)

$$\begin{aligned} W_n &= (ap+(1-p)V_{n-1})W_{n-1} \\ &= (ap+(1-p)V_{n-1})(ap+(1-p)V_{n-2})W_{n-2} \\ &\;\;\vdots \\ &= \prod_{i=1}^{n-1}(ap+1-p)V_i). \end{aligned}$$

Thus, $\log W_n = \sum_{i=1}^{n-1}\log(ap+(1-p)V_i)$. By SLLN,

$$\frac{\log W_n}{n} = \frac{n-1}{n}\cdot\frac{\sum_{i=1}^{n-1}\log(ap+(1-p)V_i)}{n-1} \overset{\text{a.s.}}{\to} E\Big(\log(ap+(1-p)V_1)\Big) =: c(p).$$

(ii) $c(p) = \int \log(ap + (1-p)\nu)\mu(d\nu)$, where μ : distribution of V. By Theorem A.5.1,

$$c''(p) = -\int \left(\frac{a-\nu}{ap+(1-p)\nu} \right)^2 \mu(d\nu) < 0.$$

(iii)

$$c'(0) = \int \frac{a-\nu}{\nu}\mu(d\nu) = -1 + aE(V_1^{-1})$$
$$c'(1) = \int \frac{a-\nu}{a}\mu(d\nu) = 1 - \frac{1}{a}EV_1.$$

To guarantee that the optimal choice of p lies in (0,1), $E(V_1^{-1}) > \frac{1}{a}$ and $E(V_1) > a$ should hold.

(iv) $p = \frac{8-5a}{2a^2-10a+8}$.

3. Central Limit Theorems

3.1. **The De Moivre-Laplace Theorem.**

3.2. **Weak Convergence.**

Exercise 3.2.1. Omit.

Exercise 3.2.2. (i)
$$\begin{aligned} P\left(M_n/n^{1/\alpha} \le y\right) &= P\left(M_n \le yn^{1/\alpha}\right) \\ &= \left[P\left(X_i \le yn^{1/\alpha}\right)\right]^n \\ &= \left(1 - \left(yn^{1/\alpha}\right)^{-\alpha}\right)^n \\ &= \left[1 - \frac{y^{-\alpha}}{n}\right]^n \to \exp\left(-y^{-\alpha}\right). \end{aligned}$$

(ii)
$$\begin{aligned} P\left(n^{1/\beta}M_n \le y\right) &= P\left(M_n \le yn^{-1/\beta}\right) \\ &= \left[P\left(X_i \le yn^{-1/\beta}\right)\right]^n \\ &= \left(1 - \frac{|y|^\beta}{n}\right)^n \to \exp\left(-|y|^\beta\right). \end{aligned}$$

(iii)
$$\begin{aligned} P(M_n - \log n \le y) &= P(M_n \le y + \log n) \\ &= [P(X_i \le y + \log n)]^n \\ &= \left(1 - e^{-y-\log n}\right)^n \\ &= \left(1 - \frac{e^{-y}}{n}\right)^n \to \exp\left(-e^{-y}\right). \end{aligned}$$

Exercise 3.2.3. (i)
$$\begin{aligned}\int_x^\infty e^{-\frac{y^2}{2}}\,dy &= \left[-\frac{1}{y}e^{-\frac{y^2}{2}}\right]_x^\infty - \int_x^\infty \frac{1}{y^2}e^{-\frac{y^2}{2}}\,dy \\ &= \left[-\frac{1}{y}e^{-\frac{y^2}{2}}\right]_x^\infty - \left[\left[-\frac{1}{y^3}e^{-\frac{y^2}{2}}\right]_x^\infty - \int_x^\infty \frac{3}{y^4}e^{-\frac{y^2}{2}}\,dy\right].\end{aligned}$$

The first equality gives us
$$\int_x^\infty e^{-\frac{y^2}{2}}\,dy \le \frac{1}{x}e^{-\frac{x^2}{2}}.$$

The second equality gives us
$$\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-\frac{x^2}{2}} \le \int_x^\infty e^{-\frac{y^2}{2}}\,dy.$$

Thus, $P(X_i > x) \sim \frac{1}{x}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$. Then,
$$\frac{P(X_i > x+\theta/x)}{P(X_i > x)} \sim \frac{\frac{1}{\sqrt{2\pi}(x+\frac{\theta}{x})}\exp\left(-\frac{1}{2}\left(x+\frac{\theta}{x}\right)^2\right)}{\frac{1}{\sqrt{2\pi}x}\exp\left(-\frac{1}{2}x^2\right)} \to e^{-\theta}$$

(ii) It's clear that $b_n \to \infty$ as $n \to \infty$. By (i), $P\left(X_i > b_n + \frac{x}{b_n}\right)\cdot n \to e^{-x}$.
$$\begin{aligned}P(b_n(M_n-b_n)\le x) &= P\left(M_n \le b_n + \frac{x}{b_n}\right) \\ &= \left[P\left(X_i \le b_n+\frac{x}{b_n}\right)\right]^n \\ &= \left[1-P\left(X_i > b_n+\frac{x}{b_n}\right)\right]^n \\ &= \left[1-P(X_i>b_n)\cdot\frac{P\left(X_i>b_n+\frac{x}{b_n}\right)}{P\left(X_i>b_n\right)}\right]^n \\ &= \left[1-\frac{1}{n}\cdot\frac{P\left(X_i>b_n+\frac{x}{b_n}\right)}{P\left(X_i>b_n\right)}\right]^n \\ &\to \exp\left(-e^{-x}\right).\end{aligned}$$

(iii) **Claim.** $b_n \sim \sqrt{2\log n}$.

Proof. ($\le$) By (i), $P(X_i > \sqrt{2\log n}) \sim \frac{1}{n\sqrt{4\pi\log n}}$. Thus, for large n, $b_n \le \sqrt{2\log n}$.
($\ge$) Similar argument with $\sqrt{2\log n - 2\log\log n}$

□

Now, by (iii), $P(b_n(M_n-b_n)\le x) \to \exp\left(-e^{-x}\right)$. Let x_n and y_n be $x_n\to\infty$ and $y_n \to -\infty$. Then,
$$P\left(M_n - b_n \le \frac{x_n}{b_n}\right) \to 1$$

$$
\begin{aligned}
&P\left(M_n - b_n \le \frac{y_n}{b_n}\right) \to 0 \\
\Rightarrow &P\left(\frac{y_n}{b_n} \le M_n - b_n \le \frac{x_n}{b_n}\right) \to 1 \\
\Rightarrow &P\left(\frac{y_n}{b_n\sqrt{2\log n}} \le \frac{M_n}{\sqrt{2\log n}} - \frac{b_n}{\sqrt{2\log n}} \le \frac{x_n}{b_n\sqrt{2\log n}}\right) \to 1.
\end{aligned}
$$

By choosing x_n, y_n with $o\left(b_n^2\right)$, we can show that

$$\frac{M_n}{\sqrt{2\log n}} \xrightarrow{P} 1.$$

Exercise 3.2.4. By Skorohod's Representation Theorem(Thm 3.2.8 in Textbook), there exists Y_n and Y such that $Y_n \overset{d}{=} X_n$ and $Y \overset{d}{=} X$ with $Y_n \xrightarrow{a.s.} Y$. By Continuous Mapping Theorem (Exercise 1.3.3), $Eg(Y_n) \xrightarrow{a.s.} Eg(Y)$. By Fatou's Lemma, we get

$$\liminf_n Eg(X_n) = \liminf_n Eg(Y_n) \ge Eg(Y) = Eg(X).$$

Exercise 3.2.5. By Skorohod's Representation Theorem, there exists Y_n and Y with distribution F_n and F, respectively, such that $Y_n \xrightarrow{a.s.} Y$. By Theorem 1.6.8 in Textbook,

$$\int h(x)dF_n(x) \to \in h(x)dF(x).$$

Exercise 3.2.6. Define $L(F,G) = \{c : F(x-c) - c \le G(x) \le F(x+c) + c \text{ for all } x\}$.
Since any distribution is non-decreasing,

(a) if $a \in L(F,G)$, then for any $b \ge a$ are also in $L(F,G)$

(b) $L(F,G)$ contains only non-negative constants.

(1) ρ is a metric.

(i) $\rho(F,G) = 0 \iff F = G$.

($\Leftarrow$) Trivial.

($\Rightarrow$) For any $\epsilon > 0$, $F(x-\epsilon) - \epsilon \le G(x) \le F(x+\epsilon) + \epsilon$ for all x. Letting $\epsilon \searrow 0$, we get $F(x^-) \le G(x) \le F(x)$ for all x. Thus, for $x \in C(F)$, $F(x) = G(x)$. For $x \in D(F)$, since $D(F)$ is at most countable(Exercise 1.2.3), we can find $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in C(F)$ with $x_n \downarrow x$. Since $F(x_n) = G(x_n)$, $F(x) = \lim_n F(x_n) = \lim_n G(x_n) = G(x)$. Therefore, $F = G$

(ii) $\rho(F,G) = \rho(G,F)$.

Suppose $a \in L(F,G)$. Then, for all x,

$$
\begin{aligned}
F(x) - a \le G(x+a) \le F(x+2a) + a \\
F(x-2a) - a \le G(x-a) \le F(x) + a.
\end{aligned}
$$

Thus, $G(x-a) - a \le F(x) \le G(x+a) + a$ for all x. That is, $a \in L(G,F)$. Therefore, $\rho(G,F) \le \rho(F,G)$. By symmetry, it also holds that $\rho(F,G) \le \rho(G,F)$.

(iii) $\rho(F,H) \le \rho(F,G) + \rho(G,H)$.

Suppose $a \in L(F,G)$ and $b \in L(G,H)$. Then, for all x,

$$F(x-a-b) - a - b \le G(x-b) - b \le H(x) \le G(x+b) + b \le F(x+a+b) + a + b.$$

That is, $a+b \in L(F,H)$. Thus, $\rho(F,H) \le \rho(F,G)+\rho(G,H)$.

By (i)-(iii), ρ is a metric.

(2) $\rho(F_n, F) \to 0$ if and only if $F_n \Rightarrow F$.

($\Rightarrow$) Since $\rho(F_n, F) \to 0$, for given $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that $\rho(F_n, F) < \epsilon$ for $n > N_\epsilon$. By (a), $\epsilon \in L(F_n, F)$ for $n > N_\epsilon$. That is, if $n > N_\epsilon$,

$$F(x-\epsilon) - \epsilon \le F_n(x) \le F(x+\epsilon) + \epsilon, \quad \forall x.$$

Letting $n \to \infty$,

$$F(x-\epsilon) - \epsilon \le \liminf F_n(x) \le \limsup F_n(x) \le F(x+\epsilon) + \epsilon, \quad \forall x.$$

For $x \in C(F)$, letting $\epsilon \searrow 0$,

$$F(x) \le \liminf F_n(x) \le \limsup F_n(x) \le F(x).$$

Thus, $F_n(x) \to F(x)$.

($\Leftarrow$) Since $F_n \Rightarrow F$, $\lim_{n\to\infty} F_n(x) = F(x)$ for $x \in C(F)$. Fix $\epsilon > 0$. Since $D(F)$ is at most countable, we can find $\{x_1, \ldots, x_{k_\epsilon}\} \in C(F)$ such that

$$F(x_1) < \epsilon, \ F\left(x_{k_\epsilon}\right) > 1-\epsilon \ \text{ and } \ |x_i - x_{i-1}| < \epsilon \text{ for } i = 2, \ldots, k_\epsilon.$$

For each x_i, there exists $N_i(\epsilon) \in \mathbb{N}$ such that $F(x_i) - \epsilon < F_n(x_i) < F(x_i) + \epsilon$ if $n > N_i(\epsilon)$. Let $N(\epsilon) = \max\{N_1(\epsilon), \ldots, N_{k_\epsilon}(\epsilon)\}$. Then, for $n > N(\epsilon)$,

(i) if $x < x_1$,

$$F(x-2\epsilon) - 2\epsilon \le F(x_1) - 2\epsilon < 0 \le F_n(x) \le F_n(x_1) < F(x_1) + \epsilon < 2\epsilon \le F(x+2\epsilon) + 2\epsilon.$$

(ii) if $x > x_{k_\epsilon}$,

$$F(x-2\epsilon) - 2\epsilon \le 1 - 2\epsilon < F\left(x_{k_\epsilon}\right) - \epsilon < F_n\left(x_{k_\epsilon}\right) \le F_n(x) \le 1 < F\left(x_{k_\epsilon}\right) + 2\epsilon \le F(x+2\epsilon) + 2\epsilon.$$

(iii) if $x_{i-1} \le x \le x_{i+1}$ for $i = 2, \ldots, k_\epsilon$,

$$F(x-\epsilon) - \epsilon \le F(x_{i-1}) - \epsilon < F_n(x_{i-1}) \le F_n(x) \le F_n(x_i) < F(x_i) + \epsilon \le F(x+\epsilon) + \epsilon.$$

By (i)-(iii), $2\epsilon \in L(F, F_n)$ for $n > N(\epsilon)$. Since $\epsilon > 0$ is arbitrarily chosen, we can conclude that $\rho(F_n, F) \to 0$.

Exercise 3.2.7. $X \sim F$ and $Y \sim G$. Recall $L(F,G)$ defined in Exercise 3.2.6.
Define $K(X,Y) = \{\epsilon \ge 0 : P(|X-Y| > \epsilon) \le \epsilon\}$. If one can show that $K(X,Y) \subset L(F,G)$, then it's done. Suppose $a \in K(X,Y)$. Then, $P(|X-Y| > a) \le a$. Since

$$P(X \le x-a) - P(|X-Y| > a) \le P(Y \le x) \le P(X \le x+a) + P(|X-Y| > a) \ \forall x,$$

we get

$$P(X \le x-a) - a \le P(Y \le x) \le P(X \le x+a) + a, \ \forall x.$$

Thus, $a \in L(F,G)$

Exercise 3.2.8. Recall $K(X,Y)$ defined in Exercise 3.2.7. It is easy to check that if $a \in K(X,Y)$, then any $b > a$ is also in $K(X,Y)$.
If $a = 0$, then $X = Y$, and thus $\beta(X,Y) = 0$. Now, suppose $a > 0$.

(1) $\frac{a^2}{1+a} \leq \beta(X,Y)$.
Fix $0 < \epsilon < a$. Since $\epsilon \notin K(X,Y)$, $\epsilon < P(|X-Y| > \epsilon)$. By Exercise 2.3.6 (b) ($\Rightarrow$), we get

$$\epsilon < P(|X-Y| > \epsilon) \leq \frac{1+\epsilon}{\epsilon} E\left(\frac{|X-Y|}{1+|X-Y|}\right).$$

Thus, $\frac{\epsilon^2}{1+\epsilon} < \beta(X,Y)$. Letting $\epsilon \nearrow a$, we get

$$\frac{a^2}{1+a^2} \leq \beta(X,Y).$$

(2) $\beta(X,Y) \leq \frac{2a}{1+a}$.
For all $b > a$,

$$\begin{aligned}
\beta(X,Y) &= \int \frac{|X-Y|}{1+|X-Y|} dP \\
&= \int_{|X-Y|>b} \frac{|X-Y|}{1+|X-Y|} dP + \int_{|X-Y|\leq b} \frac{|X-Y|}{1+|X-Y|} dP \\
&\leq \int_{|X-Y|>b} 1 dP + \int_{|X-Y|\leq b} \frac{b}{1+b} dP \\
&= P(|X-Y| > b) + \frac{b}{1+b} \cdot P(|X-Y| \leq b) \\
&= P(|X-Y| > b) \cdot \left(1 - \frac{b}{1+b}\right) + \frac{b}{1+b} \\
&\leq \frac{b}{1+b} + \frac{b}{1+b} = \frac{2b}{1+b}.
\end{aligned}$$

The last inequality holds since $b \in K(X,Y)$. Letting $b \searrow a$, we get $\beta(X,Y) \leq \frac{2a}{1+a}$.

Exercise 3.2.9. Since $F_n \Rightarrow F$ and F is continuous, $F_n(x) \to F(x)$, $\forall x$. Fix $k \in \mathbb{N}$. Pick $x_{j,k} = F^{-1}\left(\frac{j}{k}\right)$ for $j = 1, \ldots, k-1$ with $x_{0,k} = -\infty$ and $x_{k,k} = \infty$. For each j, there exists $N_k(j) \in \mathbb{N}$ such that $|F_n(x_{j,k}) - F(x_{j,k})| < \frac{1}{k}$ if $n > N_k(j)$. Let $N_k = \max\{N_k(1), \ldots, N_k(k-1)\}$. Then, for $n > N_k$, if $x_{j-1,k} < x < x_{j,k}$, then

$$F(x) - \frac{2}{k} \leq F(x_{j-1,k}) - \frac{1}{k} < F_n(x_{j-1,k}) \leq F_n(x) \leq F_n(x_{j,k}) < F(x_{j,k}) + \frac{1}{k} \leq F(x) + \frac{2}{k}.$$

That is, $|F_n(x) - F(x)| < \frac{2}{k}$. Thus, $\sup_x |F_n(x) - F(x)| \to 0$ as $n \to \infty$.

Exercise 3.2.10. Let $X_i \overset{iid}{\sim} F$ and $F_n(\omega, x) = \frac{1}{n}\sum_{i=1}^n I(X_i(\omega) \leq x)$. By Glivenko-Cantelli Theorem(Thm 2.4.9), for each $x \in \mathbb{R}$, $F_n(\omega, x) \to F(x)$ $a.s.$ Let Ω_x be a set where $F_n(\omega, x) \to F(x)$ and let $\Omega_0 = \cap_{q\in\mathbb{Q}}\Omega_q$. Then, $P(\Omega_0) = 1$ and for each $\omega_0 \in \Omega_0$, we have $F_n(\omega, q) \to F(q)$, $\forall q \in \mathbb{Q}$. Now, for $x \in C(F)$, choose $r, s \in \mathbb{Q}$ s.t. $r < x < s$ and

$$|F(r) - F(x)| < \epsilon$$

$$|F(s) - F(x)| < \epsilon.$$

Then,

$$F_n(x) - F(x) \leq F_n(s) - F(r) \to F(s) - F(r) < 2\epsilon$$
$$F_n(x) - F(x) \geq F_n(r) - F(s) \to F(r) - F(s) > -2\epsilon.$$

Therefore, $-2\epsilon \leq \liminf(F_n(x) - F(x)) \leq \limsup(F_n(x) - F(x)) \leq 2\epsilon$.

Exercise 3.2.11. Since X_n, $1 \leq n \leq \infty$, are integer valued, $\mathbb{R}\backslash\mathbb{Z} \subset C(F_n)$, where F_n is df of X_n.

($\Rightarrow$) For $1 \leq n < \infty$,

(i) $m \in \mathbb{R}\backslash\mathbb{Z}$

$P(X_n = m) = 0 = P(X_\infty = m)$.

(ii) $m \in \mathbb{Z}$

$$\begin{aligned} P(X_n = m) &= F_n\left(m + \frac{1}{2}\right) - F_n\left(m - \frac{1}{2}\right) \\ &\to F_\infty\left(m + \frac{1}{2}\right) - F_\infty\left(m - \frac{1}{2}\right) \\ &= P(X_\infty = m). \end{aligned}$$

By (i), (ii), $P(X_n = m) \to P(X_\infty = m)$, $\forall m$.

($\Leftarrow$) Want to show that $F_n(x) \to F_\infty(x)$, $\forall x \in C(F_\infty)$.

Fix $\epsilon > 0$. Since F_∞ can have positive probability only on $\mathbb{Z}$, which is countable, we can pick $K(\epsilon) := \{k_1, \cdots, k_{l_\epsilon}\} \subset \mathbb{Z}$ such that

$$\sum_{i=1}^{l_\epsilon} P(X_\infty = k_i) > 1 - \epsilon. \tag{11a}$$

By assumption, for each k_i, there exists $N_\epsilon(i) \in \mathbb{N}$ such that for $n > N_\epsilon(i)$,

$$|P(X_n = k_i) - P(X_\infty = k_i)| < \frac{\epsilon}{l_\epsilon}.$$

Let $N_\epsilon = \max\{N_\epsilon(1), \cdots, N_\epsilon(l_\epsilon)\}$. Then, for $n < N_\epsilon$,

$$\left|\sum_{i=1}^{l_\epsilon}[P(X_n = k_i) - P(X_\infty = k_i)]\right| \leq \sum_{i=1}^{l_\epsilon}|P(X_n = k_i) - P(X_\infty = k_i)| =: P_0 < \epsilon$$

Also, by (11a), we have

$$P_1 := \sum_{k \in \mathbb{Z}\backslash K(\epsilon)} P(X_\infty = k) < \epsilon.$$

Thus,

$$P_2 := \left|\sum_{k \in \mathbb{Z}\backslash K(\epsilon)} P(X_n = k)\right|$$

$$
\begin{aligned}
&= \left|1 - \sum_{i=1}^{l_\epsilon} P(X_n = k_i)\right| \\
&= \left|1 - \sum_{i=1}^{l_\epsilon} P(X_\infty = k_i) + \sum_{i=1}^{l_\epsilon} P(X_\infty = k_i) - \sum_{i=1}^{l_\epsilon} P(X_n = k_i)\right| \\
&\le \left|\sum_{k\in\mathbb{Z}\setminus K(\epsilon)} P(X_\infty = k)\right| + \left|\sum_{i=1}^{l_\epsilon} \big[P(X_n = k_i) - P(X_\infty = k_i)\big]\right| \\
&< 2\epsilon.
\end{aligned}
$$

Now, for $x \in C(F_\infty)$, let $Z_x = \mathbb{Z} \cap (-\infty, x]$.

$$
\begin{aligned}
F_n(x) &= \sum_{k\in Z_x} P(X_n = k) \\
&= \sum_{k\in Z_x\cap K(\epsilon)} P(X_n = k) + \sum_{k\in Z_x\setminus K(\epsilon)} P(X_n = k).
\end{aligned}
$$

Similarly,

$$
F_\infty(x) = \sum_{k\in Z_x\cap K(\epsilon)} P(X_\infty = k) + \sum_{k\in Z_x\setminus K(\epsilon)} P(X_\infty = k).
$$

For $n > N_\epsilon$,

$$
\begin{aligned}
|F_n(x) - F_\infty(x)| &= \left|\sum_{k\in Z_x\cap K(\epsilon)} \big[P(X_n = k) - P(X_\infty = k)\big] + \sum_{k\in Z_x\setminus K(\epsilon)} \big[P(X_n = k) - P(X_\infty = k)\big]\right| \\
&\le \left|\sum_{k\in Z_x\cap K(\epsilon)} \big[P(X_n = k) - P(X_\infty = k)\big]\right| + \left|\sum_{k\in Z_x\setminus K(\epsilon)} P(X_n = k)\right| + \left|\sum_{k\in Z_x\setminus K(\epsilon)} P(X_\infty = k)\right| \\
&\le \sum_{k\in Z_x\cap K(\epsilon)} |P(X_n = k) - P(X_\infty = k)| + \left|\sum_{k\in Z_x\setminus K(\epsilon)} P(X_n = k)\right| + \left|\sum_{k\in Z_x\setminus K(\epsilon)} P(X_\infty = k)\right|.
\end{aligned}
$$

Note that $\begin{cases} Z_x \cap K(\epsilon) \subset K(\epsilon) = \{k_1, \cdots, k_{l_\epsilon}\} \\ Z_x\setminus K(\epsilon) \subset \mathbb{Z}\setminus K(\epsilon) \end{cases}$. Thus,

$$
|F_n(x) - F_\infty(x)| \le P_0 + P_1 + P_2 < 4\epsilon.
$$

Therefore, $F_n(x) \to F_\infty(x)$, and thus $X_n \Rightarrow X_\infty$.

Exercise 3.2.12. (1) If $X_n \to X$ in probability, then $X_n \Rightarrow X$.

Since $X_n \xrightarrow{p} X$, for given $\epsilon > 0$, $P(|X_n - X| > \epsilon) \to 0$.

(i) $P(X_n \le x) = P(X_n \le x, |X_n - X| \le \epsilon) + P(X_n \le x, |X_n - x| > \epsilon)$
$\le P(X \le x + \epsilon) + P(|X_n - X| > \epsilon)$.

Then, $\limsup_n P(X_n \le x) \le P(X \le x + \epsilon)$. By letting $\epsilon \searrow 0$, we get

$$
\limsup_n P(X_n \le x) \le P(X \le x).
$$

(ii) $P(X_n > x) = P(X_n > x, |X_n - X| \leq \epsilon) + P(X_n > x, |X_n - X| > \epsilon)$
$\leq P(X > x - \epsilon) + P(|X_n - X| > \epsilon)$.

Then,

$$\begin{aligned}
&\limsup_n P(X_n > x) \leq P(X > x - \epsilon) \\
&\Rightarrow \limsup_n \,[1 - P(X_n \leq x)] \leq 1 - P(X \leq x - \epsilon) \\
&\Rightarrow \liminf_n P(X_n \leq x) \geq P(X \leq x - \epsilon).
\end{aligned}$$

For $x \in C(F_X)$, by letting $\epsilon \searrow 0$, we get

$$\liminf_n P(X_n \leq x) \geq P(X \leq x).$$

By (i), (ii), for $x \in C(F_X)$,

$$F_n(x) = P(X_n \leq x) \to P(X \leq x) = F(x).$$

Thus, $X_n \Rightarrow X$.

(2) Since $X_n \Rightarrow c$, $P(X_n \leq y) \to \begin{cases} 0, & y < c \\ 1, & y > c \end{cases}$. Thus, for any $\epsilon > 0$,

$$\begin{aligned}
P(|X_n - c| > \epsilon) &= P(X_n > c + \epsilon) + P(X_n < c - \epsilon) \\
&= 1 - P(X_n \leq c + \epsilon) + P(X_n < c - \epsilon) \\
&\to 0.
\end{aligned}$$

Exercise 3.2.13. Since $\begin{cases} X_n \Rightarrow X, \;\; P(X_n \leq x) \to P(X \leq x), \;\; \forall x \in C(F_X). \\ Y_n \xrightarrow{P} c, \;\; P(|Y_n - c| > \epsilon) \to 0, \;\; \forall \epsilon > 0. \end{cases}$

(i) $P(X_n + Y_n \leq z) = P(X_n + Y_n \leq z, |Y_n - c| \leq \epsilon) + P(X_n + Y_n \leq z, |Y_n - c| > \epsilon)$
$\geq P(X_n \leq z - c + \epsilon) + P(|Y_n - c| > \epsilon)$.

Since $D(F_X)$ is at most countable, we may assume that $z - c + \epsilon \in C(F_X)$. Then, $\limsup_n P(X_n + Y_n \leq z) \leq P(X \leq z - c + \epsilon)$. Letting $\epsilon \searrow 0$,

$$\limsup_n P(X_n + Y_n \leq z) \leq P(X \leq z - c) = P(X + c \leq z).$$

(ii) $P(X_n + Y_n > z) = P(X_n + Y_n > z, |Y_n - c| \leq \epsilon) + P(X_n + Y_n > z, |Y_n - c| > \epsilon)$
$\leq P(X_n > z - c - \epsilon) + P(|Y_n - c| > \epsilon)$.

By similar argument, we get

$$\liminf_n P(X_n + Y_n \leq z) \geq P(X \leq z - c - \epsilon).$$

For $z - c \in C(F_X)$, by letting $\epsilon \to 0$,

$$\liminf_n P(X_n + Y_n \leq z) \geq P(X \leq z - c) = P(X + c \leq z).$$

Note that $z \in C(F_{X+c})$ if and only if $z - c \in C(F_X)$. By (i), (ii), we get

$$P(X_n + Y_n \le z) \to P(X + c \le z).$$

Exercise 3.2.14. Since $\begin{cases} X_n \Rightarrow X, \;\; P(X_n \le x) \to P(X \le x), \;\; \forall x \in C(F_X). \\ Y_n \xrightarrow{P} c, \;\; P(|Y_n - c| > \epsilon) \to 0, \;\; \forall \epsilon > 0. \end{cases}$
For simple proof, let's assume $0 < \epsilon < c$.

(i) $P(X_nY_n \le z) = P(X_nY_n \le z, |Y_n - c| \le \epsilon) + P(X_nY_n \le z, |Y_n - c| > \epsilon)$

$$\le P\left(X_n \le \frac{z}{c-\epsilon}\right) + P(|Y_n - c| > \epsilon).$$

Let $\frac{z}{c-\epsilon} \in C(F_X)$. Then,

$$\limsup_n P(X_nY_n \le z) \le P\left(X \le \frac{z}{c-\epsilon}\right).$$

Letting $\epsilon \to 0$,

$$\limsup_n P(X_nY_n \le z) \le P\left(X \le \frac{z}{c}\right) = P(cX \le z).$$

(ii) $P(X_nY_n > z) = P(X_nY_n > z, |Y_n - c| \le \epsilon) + P(X_nY_n > z, |Y_n - c| > \epsilon)$ Similarly,

$$\le P\left(X_n > \frac{z}{c+\epsilon}\right) + P(|Y_n - c| > \epsilon).$$

$$\liminf_n P(X_nY_n \le z) \ge P\left(X \le \frac{z}{c+\epsilon}\right).$$

For $\frac{z}{c} \in C(F_X)$, by letting $\epsilon \to 0$,

$$\liminf_n P(X_nY_n \le z) \ge P\left(X \le \frac{z}{c}\right) = P(cX \le z).$$

Note that $z \in C(F_{cX})$ if and only if $\frac{z}{c} \in C(F_X)$. By (i),(ii), we get

$$P(X_nY_n \le z) \to P(cX \le z).$$

Exercise 3.2.15. Step 1. $X_n^i \overset{d}{\equiv} Y_i / (\sum_{m=1}^n Y_m^2/n)^{1/2}$.
Note that uniformly distributed over the surface of the unit sphere is the only distribution which is invariant to rotation and belongs to unit sphere.
Since $Y_i \overset{iid}{\sim} N(0,1)$, $Y = (Y_1, \cdots, Y_n) \sim N(\mathbf{0}, I_n)$. For any orthogonal matrix A, $AY \sim N(\mathbf{0}, I_n) \overset{d}{\equiv} Y$. Then, $\frac{Y}{||Y||_2} \overset{d}{\equiv} \frac{AY}{||AY||_2}$. Thus, $\frac{Y}{||Y||_2} = (Y_1/\sum_{m=1}^n Y_m^2, \cdots, Y_m/\sum_{m=1}^n Y_m^2)$ is uniformly distributed over the surface of the unit sphere. Then, $X_n \overset{d}{\equiv} \sqrt{n}\frac{Y}{||Y||_2}$, and thus

$$X_n^i \overset{d}{\equiv} Y_i \Bigg/ \left(\sum_{m=1}^n Y_m^2/n\right)^{1/2}.$$

Step 2. $Y_i/(\sum_{m=1}^n Y_m^2/n)^{1/2} \Rightarrow N(0,1)$.
For $Y_i \sim N(0,1)$, $EY_i^2 = 1 < \infty$. Thus, by WLLN,

$$\sum_{m=1}^{n} Y_m^2 \Big/ n \xrightarrow{P} 1.$$

By Exercise 3.2.12(1), $\sum_{m=1}^n Y_m^2/n \Rightarrow 1$. Note that $f(x) = \frac{1}{\sqrt{x}} I_{(x>0)}$ is continuous on $\mathbb{R}\backslash\{0\}$ and $P(\sum_{m=1}^n Y_m^2/n = 0) = P(Y_1 = \cdots = Y_n = 0) = 0$. By Continuous Mapping Theorem (Theorem 3.2.10), we get

$$\left(\sum_{m=1}^{n} Y_m^2 \Big/ n\right)^{-1/2} \Rightarrow 1.$$

By Exercise 3.2.14,

$$Y_i \cdot \left(\sum_{m=1}^{n} Y_m^2 \Big/ n\right)^{-1/2} \Rightarrow N(0,1).$$

Thus, $X_n^1 \Rightarrow$ standard normal.

Exercise 3.2.16. For $1 < M < \infty$, since $EY_n^\beta \to 1$, there exists $N \in \mathbb{N}$ such that for $n > N$, $EY_n^\beta \le M$. Therefore, we can drop finitely many Y_n's and say $\{F_n\}$ is tight. (with $\varphi(y) = |y|^\beta$)
For every subsequence (F_{n_k}), by Helly and Tightness Theorem, there exists $(F_{n(m_k)})$ such that $F_{n(m_k)} \Rightarrow F$, where F is a df. Let Y be a r.v. with F.
For $g(y) = |y|^\beta$ and $h(y) = |y|^\alpha$, we have

(0) g, h are continuous.
(i) $g(y) > 0$
(ii) $|h(y)|/g(y) = |y|^{\alpha-\beta} \to 0$ as $|y| \to \infty$
(iii) $F_{n(m_k)} \Rightarrow F$
(iv) $\int |y|^\beta dF_{n(m_k)}(y) = EY^\beta_{n(m_k)} \le M < \infty$.

By Exercise 3.2.5,

$$\int |y|^\alpha dF_{n(m_k)}(y) = EY^\alpha_{n(m_k)} \to EY^\alpha = \int |y|^\alpha dF(y).$$

Since convergent real sequence has the unique limit, $EY^\alpha = 1$. Also, similar argument for $h(y) = |y|^\gamma$ with $\gamma \in (\alpha, \beta)$ gives us that

$$EY^\gamma_{n(m_k)} \to EY^\gamma.$$

From Jensen's Inequality with following functions, we get $EY_n^\gamma \to 1$.

(1) $\varphi_1(y) = |y|^{\frac{\gamma}{\alpha}}$.

$$(EY_n^\alpha)^{\frac{\gamma}{\alpha}} \le E\left(Y_n^\alpha\right)^{\frac{\gamma}{\alpha}} = EY_n^\gamma \quad \Rightarrow \quad 1 \le \liminf_n EY_n^\gamma.$$

(2) $\varphi_2(y) = |y|^{\frac{\beta}{\gamma}}$.

$$(EY_n^\gamma)^{\frac{\beta}{\gamma}} \le E\left(Y_n^\gamma\right)^{\frac{\beta}{\gamma}} = EY_n^\gamma \quad \Rightarrow \quad \limsup_n EY_n^\gamma \le 1.$$

Therefore, $EY^\gamma = 1$.
For strictly convex $\varphi(y) = |y|^{\frac{\gamma}{\alpha}}$, we have $(EY^\alpha)^{\frac{\gamma}{\alpha}} = 1 = E(Y^\alpha)^{\frac{\gamma}{\alpha}} = EY^\gamma$. By Exercise 1.6.1, $Y^\alpha = 1$ a.s. Since $\alpha > 0$, $Y = 1$ a.s. Since $Y_{n(m_k)} \xrightarrow{P} 1$, by Theorem 2.3.2,

$$Y_n \xrightarrow{P} 1.$$

Exercise 3.2.17. Suppose not. Then, for some $K < \infty$ and $y < 1$, there exists a r.v. X with $EX^2 = 1$ and $EX^4 \le K$ such that $P(|X| > y) < c$ for all $c > 0$. With $c = \frac{1}{n}$, $n \in \mathbb{N}$, we can construct a sequence (X_n) such that $P(|X_n| > y) \to 0$, $EX_n^2 = 1$ and $EX_n^4 \le K$. Theorem 3.2.14 with $\varphi(x) = x^2$ gives us that $\{X_n\}$ is tight. Then, by Helly and Tightness Theorem, there exists (X_{n_m}) such that $X_{n_m} \Rightarrow X_\infty$. By Generalized DCT with $h(x) = x^2$ and $g(x) = x^4$, we get $EX_\infty^2 = \lim_{m\to\infty} EX_{n_m}^2 = 1$. However, since $P(|X_n| > y) \to 0$, $P(|X_n| \le y) \to 1$. Then, $EX_\infty^2 \le y^2 < 1$, which is a contradiction.

3.3. **Characteristic Functions.**

Exercise 3.3.1. Let X be a r.v. with ch.f φ and X' be an independent copy of $-X$. Then, its ch.f $\varphi'(t) = Ee^{it(-X)} = \overline{\varphi(t)}$. By Theorem 3.3.2 and Lemma 3.3.9 in Textbook, Re φ is ch.f of $\frac{1}{2}(X + X')$ and $|\varphi|^2$ is ch.f of $X + X'$.

Exercise 3.3.2. (i) Let

$$\begin{aligned} I_T &= \frac{1}{2T}\int_{-T}^{T} e^{-ita}\varphi(t)\,dt \\ &= \frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{\infty} e^{-ita}e^{itx}\,\mu(dx)\,dt \\ &= \frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{\infty} e^{it(x-a)}\,\mu(dx)\,dt \end{aligned}$$

Since $\left|e^{it(x-a)}\right| \le 1$ and $[-T, T]$ is a finite interval, by Fubini,

$$I_T = \frac{1}{2T}\int_{-\infty}^{\infty}\int_{-T}^{T} \cos(t(x-a)) + i\sin(t(x-a))\,dt\,\mu(dx).$$

Since $\sin t$ is symmetric about the origin,

$$I_T = \frac{1}{2T}\int_{-\infty}^{\infty}\int_{-T}^{T} \cos(t(x-a))\,dt\,\mu(dx)$$

Since $\left|\int_0^x \cos(y)dy\right| = |\sin(y)| \le |x|$, $\left|\frac{1}{2T}\int_{-T}^{T}\cos(t(x-a))\,dt\right| \le 1$: integrable. By DCT,

$$\lim_{T\to\infty} I_T = \int_{-\infty}^{\infty}\left[\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\cos(t(x-a))\,dt\right]\mu(dx).$$

(1) $x \neq a$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \cos(t(x-a))\, dt = \lim_{T\to\infty} \frac{\sin((x-a)T)}{(x-a)T} = 0.$$

(2) $x = a$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \cos(t(x-a))\, dt = 1.$$

By (1) and (2), we get

$$\lim_{T\to\infty} I_T = \mu\big(\{a\}\big).$$

(ii) Let μ be the distribution of X.

$$\begin{aligned}
\varphi\left(\frac{2\pi}{h}+t\right) &= \int e^{i(\frac{2\pi}{h}+t)x}\mu(dx) \\
&= \int_{h\mathbb{Z}} e^{i(\frac{2\pi}{h}+t)x}\mu(dx) \\
&= \int_{h\mathbb{Z}} e^{itx}\mu(dx) \\
&= \int e^{itx}\mu(dx) = \varphi(t). \qquad \text{(2a)}
\end{aligned}$$

By (i),

$$P(X=x) = \mu(\{x\}) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{-itx}\varphi(t)dt.$$

Let $T_n = \frac{\pi}{h} + \frac{2\pi}{h}\cdot n$, with $n \geq 0$. Then, since limit of RHS above exists,

$$P(X=x) = \lim_{n\to\infty} \frac{1}{2T_n} \int_{-T_n}^{T_n} e^{-itx}\varphi(t)dt.$$

Now, for $1 \leq k \in \mathbb{N}$,

$$\begin{aligned}
\int_{T_{k-1}}^{T_k} e^{-itx}\varphi(t)dt &= \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-i(t+\frac{2k\pi}{h})x}\cdot\varphi\left(t+\frac{2k\pi}{h}\right)dt \\
&= e^{-i(\frac{2k\pi x}{h})}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-it}\varphi(t)dt
\end{aligned}$$

by (2a). Since $x \in h\mathbb{Z}$, $\frac{2k\pi x}{h} \in \mathbb{N}$. Thus, $e^{-i(\frac{2k\pi x}{h})} = 1$. Therefore,

$$\begin{aligned}
\int_{T_{k-1}}^{T_k} e^{-itx}\varphi(t)dt &= \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-it}\varphi(t)dt \\
&= \int_{-T_0}^{T_0} e^{-itx}\varphi(t)dt.
\end{aligned}$$

This also holds for $\int_{-T_k}^{-T_{k-1}} e^{-itx}\varphi(t)dt = \int_{-T_0}^{T_0} e^{-itx}\varphi(t)dt$. Then,

$$\begin{aligned}\frac{1}{2T_n}\int_{-T_n}^{T_n} e^{-itx}\varphi(t)dt &= \frac{1}{2\cdot\frac{(2n+1)}{h}\pi}\cdot(2n+1)\int_{-T_0}^{T_0} e^{-itx}\varphi(t)dt \\ &= \frac{h}{2\pi}\int_{-T_0}^{T_0} e^{-itx}\varphi(t)dt \\ &= \frac{h}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-itx}\varphi(t)dt.\end{aligned}$$

Therefore,

$$P(X=x) = \frac{h}{2\pi}\int_{-\frac{h}{\pi}}^{\frac{h}{\pi}} e^{-itx}\varphi(t)dt.$$

(iii) $X = Y + b$. If $P(X \in b + h\mathbb{Z}) = 1$, then $P(Y \in h\mathbb{Z}) = 1$. For $x \in b + h\mathbb{Z}$, $P(X = x) = P(Y = x - b)$. Then,

$$\begin{aligned}P(Y = x-b) &= \frac{h}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-it(x-b)}\varphi_Y(t)dt \\ &= \frac{h}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-itx}\cdot e^{itb}\cdot e^{-itb}\varphi_X(t)dt \\ &= \frac{h}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-itx}\varphi_X(t)dt.\end{aligned}$$

Exercise 3.3.3. Since X and Y are independent, X and $-Y$ are also independent. Then, ch.f of $X - Y$ is given by $\varphi(t)\cdot\overline{\varphi(t)} = |\varphi(t)|^2$. By Exercise 3.3.2,

$$P(X - Y = 0) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|\varphi(t)|^2dt.$$

By Exercise 2.1.5,

$$P(X-Y=0) = \sum_x \mu\big(\{x\}\big)^2.$$

Exercise 3.3.4. For exponential distribution,

$$\begin{aligned}f(y) &= e^{-y}I_{(y>0)} \\ \varphi(t) &= \frac{1}{1-it}\end{aligned}$$

And,

$$\int|\varphi(t)|dt = \int\left|\frac{1}{1-it}\right|dt = \int\frac{1}{\sqrt{1+t^2}}dt = \infty.$$

Exercise 3.3.5. ch.f of X_1 is given by

$$\begin{aligned}\varphi_1(t) &= Ee^{itX} \\ &= \int_{-1}^{1} e^{itx}\cdot\frac{1}{2}dx = \left[\frac{1}{2it}e^{itx}\right]_{-1}^{1} \\ &= \frac{1}{2it}\left(e^{it} - e^{-it}\right)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2it}(\cos t + i\sin t - \cos t + i\sin t) \\
&= \frac{\sin t}{t}.
\end{aligned}
$$

Thus, ch.f of $X_1 + \cdots + X_n$ is given by $\varphi(t) = \left(\frac{\sin t}{t}\right)^n$.
Note that

$$
\int |\varphi(t)| dt = 2\int_0^\infty \left|\frac{\sin t}{t}\right|^n dt.
$$

For $n \geq 2$,

$$
\begin{aligned}
\int_0^\infty \left|\frac{\sin t}{t}\right|^n dt &= \int_0^1 \left|\frac{\sin t}{t}\right|^n dt + \int_1^\infty \left|\frac{\sin t}{t}\right|^n dt \\
&\leq \int_0^1 1dt + \int_1^\infty \frac{1}{t^n} dt < \infty.
\end{aligned}
$$

Therefore, by Thm 3.3.14,

$$
\begin{aligned}
f(x) &= \frac{1}{2\pi}\int_{-\infty}^\infty e^{-itx} \cdot \left(\frac{\sin t}{t}\right)^n dt \\
&= \frac{1}{2\pi}\int_{-\infty}^\infty \left[\left(\frac{\sin t}{t}\right)^n \cdot \cos tx - i\left(\frac{\sin t}{t}\right)^n \cdot \sin tx\right] dt \\
&= \frac{1}{\pi}\int_0^\infty \left(\frac{\sin t}{t}\right)^n \cos tx dt.
\end{aligned}
$$

Exercise 3.3.6. In Example 3.3.16, we showed that ch.f of Cauchy distribution is given by $\exp(-|t|)$.
Since $X_1, \cdots, X_n$ are independent, by Thm 3.3.2,

$$
\varphi_{X_1+\cdots+X_n}(t) = \prod_{i=1}^n \varphi_{X_i}(t) = \prod_{i=1}^n \exp(-|t|) = \exp(-n|t|)
$$

Bu Thm 3.3.1(e),

$$
\varphi_{\frac{X_1+\cdots+X_n}{n}}(t) = \varphi_{X_1+\cdots+X_n}\left(\frac{t}{n}\right) = \exp(-|t|),
$$

which is ch.f of Cauchy. Thus, by Inversion Formula, $\frac{X_1+\cdots+X_n}{n} \overset{d}{\equiv} X_1$.

Exercise 3.3.7. $\varphi_{X_n}(t) = \exp\left(-\frac{\sigma_n^2 t^2}{2}\right)$.
Since $X_n \Rightarrow X$, $\varphi_{X_n}(t) \to \varphi_\infty(t)$, $\forall t$. For $t = 1$, $\varphi_{X_n}(1) = \exp\left(-\frac{\sigma_n^2}{2}\right)$. Thus,

$$
\begin{aligned}
\varphi_{X_n}(1) \text{ converges } &\iff \log\varphi_{X_n}(1) \text{ converges} \\
&\iff -\frac{\sigma_n^2}{2} \text{ converges.}
\end{aligned}
$$

Therefore, there exists σ^2 such that $\sigma_n^2 \to \sigma^2$. Since $\sigma_n^2 \in (0, \infty)$, we can say that $\sigma^2 \in [0, \infty]$.
Now, suppose $\sigma^2 = \infty$. Then $\varphi_{X_n}(t) = \exp\left(-\frac{\sigma_n^2 t^2}{2}\right) \to 0$ if $t \neq 0$. However, $\varphi_{X_n}(0) = 1$, $\forall n$.
Thus, continuity of φ_X at 0 is violated, which means that $X_n \not\Rightarrow X$. Thus, $\sigma_n^2 \to \sigma^2 \in [0, \infty)$.

Exercise 3.3.8. Let $\varphi_n(t)$ be the ch.f of X_n and $\psi_n(t)$ be the ch.f of Y_n, for $1 \le n \le \infty$. By Theorem 3.3.17(Continuity Theorem),

$$\begin{aligned}\varphi_n(t) &\to \varphi_\infty(t), \forall t\\ \psi_n(t) &\to \psi_\infty(t), \forall t.\end{aligned}$$

Since X_n and Y_n are independent, ch.f of $X_n + Y_n = \varphi_n(t)\psi_n(t)$, for $1 \le n \le \infty$. Then, $\varphi_n(t)\psi_n(t) \to \varphi_\infty(t)\psi_\infty(t)$, $\forall t$. Since $\varphi_\infty, \psi_\infty$ are ch.f, they are continuous at 0. Thus, $\varphi_\infty\psi_\infty$ is also continuous at 0. By Thm 3.3.17, $X_n + Y_n \Rightarrow X_\infty + Y_\infty$.

Exercise 3.3.9. By Thm 3.3.2, ch.f of S_n is given by

$$\varphi_{S_n}(t) = \prod_{j=1}^n \varphi_j(t).$$

Since a.s. convergence implies weak convergence, $S_n \Rightarrow S_\infty$. By Thm 3.3.17(i),

$$\varphi_{S_\infty}(t) = \lim_{n\to\infty} \varphi_{S_n}(t) = \prod_{j=1}^\infty \varphi_j(t), \ \forall t.$$

Exercise 3.3.10. By Exercise 3.3.5, $\frac{\sin t}{t}$ is ch.f of $Unif(-1,1)$. Let $Y_1, \ldots, Y_n$ be $i.i.d$ symmetric Bernoulli random variables. $\left(P(Y_1 = 1) = P(Y_1 = -1) = \frac{1}{2}\right)$. Define $X_i = \frac{Y_i}{2^i}$. Then, ch.f of $X_i = \cos\frac{t}{2^i}$, and thus ch.f of $\sum_{i=1}^\infty X_i = \prod_{i=1}^\infty \cos\frac{t}{2^i}$. **Claim.** $\sum_{i=1}^n X_i \Rightarrow Unif(-1,1)$.

$$P\left(\sum_{i=1}^n X_i = \frac{2k+1}{2^n}\right) = \frac{1}{2^n} \quad \text{for } k = -2^{n-1}, \cdots, -1, 0, 1, \cdots, 2^{n-1} - 1.$$

Then,

$$\begin{aligned}P\left(\sum_{i=1}^n X_i \le x\right) &= \max\left[\frac{1}{2^n}\left\{\left\lfloor\frac{\lfloor 2^n x\rfloor - 1}{2}\right\rfloor + 2^{n-1}\right\}, 0\right]\\ &= \max\left[\frac{1}{2} + \frac{1}{2^n}\left\lfloor\frac{\lfloor 2^n x\rfloor - 1}{2}\right\rfloor, 0\right]\end{aligned}$$

Now, for $x \in (-1,1)$, letting $n \to \infty$ gives us

$$P\left(\sum_{i=1}^\infty X_i \le x\right) = \frac{1}{2} + \frac{1}{2}x.$$

Note that for $x \in (-1,1)$, $P(Unif(-1,1) \le x) = \frac{1}{2} + \frac{1}{2}x$. Therefore, $\sum_{i=1}^n X_i \Rightarrow Unif(-1,1)$. By Theorem 3.3.17, we get the desired result.

Exercise 3.3.11. ch.f of $\frac{2X_j}{3^j} = Ee^{it\frac{2X_j}{3^j}} = \frac{1}{2} + \frac{1}{2}\cdot e^{\frac{2it}{3^j}}$. By Exercise 3.3.9, ch.f of X is given by

$$\varphi_X(t) = \prod_{j=1}^\infty \left[\frac{1}{2} + \frac{1}{2}e^{\frac{2it}{3^j}}\right].$$

Then,

$$\begin{aligned}\varphi_X\left(3^k\pi\right) &= \prod_{j=1}^{\infty}\left[\frac{1}{2}+\frac{1}{2}\exp\left(\frac{2i3^k\pi}{3^j}\right)\right] \\ &= \prod_{l=1}^{\infty}\left[\frac{1}{2}+\frac{1}{2}\exp\left(2\pi\cdot 3^{-l}i\right)\right]\cdot\prod_{j=1}^{k}\left[\frac{1}{2}+\frac{1}{2}\exp\left(2\pi\cdot 3^{k-j}i\right)\right]\end{aligned}$$

Since $3^{k-j}\in\mathbb{N}\cup\{0\}$, we get

$$\varphi_X\left(3^k\pi\right)=\prod_{j=1}^{\infty}\left[\frac{1}{2}+\frac{1}{2}\exp\left(2\pi 3^{-l}i\right)\right].$$

Exercise 3.3.12. We have $\int_{-\infty}^{\infty}|x|e^{-\frac{1}{2}x^2}dx = 2\int_0^{\infty}xe^{-\frac{1}{2}x^2}dx<\infty$. By induction, for $n\geq 2$,

$$\begin{aligned}\int_{-\infty}^{\infty}|x|^n e^{-\frac{1}{2}x^2}dx &= 2\int_0^{\infty}x^n e^{-\frac{1}{2}x^2}dx \\ &= 2\left[-x^{n-1}e^{-\frac{1}{2}x^2}\right]_0^{\infty}+2\int_0^{\infty}x^{n-1}e^{-\frac{1}{2}x^2}dx \\ &= 0+2\int_0^{\infty}x^{n-1}e^{-\frac{1}{2}x^2}dx<\infty.\end{aligned}$$

Note that

$$\begin{aligned}\varphi^{(n)}(t) &= \exp^{-\frac{1}{2}t^2} \\ &= 1-\frac{1}{2}t^2+\frac{1}{2!}\cdot\left(-\frac{1}{2}t^2\right)^2+\frac{1}{3!}\cdot\left(-\frac{1}{2}t^2\right)^3+\cdots \\ &= 1+\sum_{k=1}^{\infty}\frac{1}{k!\cdot(-2)^k}t^{2k}.\end{aligned}$$

By Thm 3.3.18, we have $i^nEX^n=\varphi^{(n)}(t)|_{t=0}$. Therefore,

$$EX^{2n}=\frac{(2n)!}{2^n n!}.$$

Exercise 3.3.13. (i) Fix $\epsilon>0$. For $0<h<\frac{\epsilon^2}{4}$,

$$\begin{aligned}|\varphi_i(t+h)-\varphi_i(t)| &\leq E|e^{ihX_i}-1| \\ &\leq E\min(|hX_i|,2) \\ &\leq E\left(|hX_i|;|X_i|\leq\frac{2}{\sqrt{h}}\right)+E\left(2;|X_i|>\frac{2}{\sqrt{h}}\right) \\ &\leq 2\sqrt{h}+2P\left(|X_i|>\frac{2}{\sqrt{h}}\right) \\ &\leq 2\sqrt{h}+\epsilon \\ &\leq 2\epsilon.\end{aligned}$$

The first inequality comes from Theorem 3.3.1(d). The second inequality comes from (3.3.3). The fifth inequality comes from tightness.

(ii) Since $\mu_n \Rightarrow \mu_\infty$, $\{\mu_n\}_{n\in\mathbb{N}}$ is tight, and thus $\{\mu_n\}_{1\le n\le\infty}$ is tight.
Then, by (i), for $\epsilon > 0$, choose $\delta > 0$ so that

$$\text{if } |h| < \delta, \text{ then } |\varphi_n(t+h) - \varphi_n(t)| < \epsilon, \text{ for } 1 \le n \le \infty.$$

Without loss of generality, assume the compact set is $[-K, K]$.
Choose integer $m > \frac{1}{\delta}$ and let $x_j = \frac{j}{m}\cdot K$ for $-m \le j \le m$. By Continuity Theorem, we have $\varphi_n(t) \to \varphi_\infty(t)$ for all t. Thus, for x_j, there exists $n_j \in \mathbb{N}$ such that if $n > n_j$, then $|\varphi_n(x_j) - \varphi_\infty(x_j)| < \epsilon$. Let $n_0 = \max\{n_{-m}, \cdots, n_m\}$. Then, for all $t \in [-K, K]$, there exists j such that $\left|t - \frac{jK}{m}\right| < \delta$. Now, for $n > n_0$,

$$\begin{aligned}
|\varphi_n(t) - \varphi_\infty(t)| &\le \left|\varphi_n(t) - \varphi_n\left(\frac{jK}{m}\right)\right| + \left|\varphi_n\left(\frac{jK}{m}\right) - \varphi_\infty\left(\frac{jK}{m}\right)\right| + \left|\varphi_\infty\left(\frac{jK}{m}\right) - \varphi_\infty(t)\right| \\
&< 3\epsilon.
\end{aligned}$$

Therefore, $\varphi_n \to \varphi_\infty$ uniformly on a compact set.
(iii) Let $X_n = \frac{1}{n}$. Then, $\varphi_n(t) = e^{\frac{it}{n}}$

Exercise 3.3.14. (i) ch.f of $\begin{cases} \frac{S_n}{n} = \varphi\left(\frac{t}{n}\right)^n \\ a = e^{iat} \end{cases}$.

Claim. $\varphi\left(\frac{t}{n}\right) \to e^{iat}$, for all t.
Proof. Since $\varphi'(0) = ia$,

$$\lim_{n\to\infty} \frac{\varphi(\frac{t}{n}) - \varphi(0)}{t/n} = ia \Rightarrow \lim_{n\to\infty} n\left(\varphi\left(\frac{t}{n}\right) - 1\right) = iat.$$

Thus, for all t,

$$\begin{aligned}
\lim_{n\to\infty} \varphi\left(\frac{t}{n}\right)^n &= \lim_{n\to\infty}\left(1 + \frac{n\left(\varphi(\frac{t}{n}) - 1\right)}{n}\right)^n \\
&= e^{iat}.
\end{aligned}$$

The second equality holds by Theorem 3.4.2. □

By *Claim* and Theorem 3.3.17, $\frac{S_n}{a} \Rightarrow a$, and thus $\frac{S_n}{a} \xrightarrow{P} a$.

(ii) In probability convergence implies weak convergence. By Theorem 3.3.17, we get the desired result.
(iii) (I think $(\varphi(h) - 1)/h \to +ia$.)
(a) Since $\log z$ is differentiable at $z = 1$,

$$\lim_{n\to\infty} \frac{\log\varphi(\frac{t}{n}) - \log 1}{\varphi(\frac{t}{n}) - 1} = \lim_{n\to\infty} \frac{\log\varphi(\frac{t}{n})}{\varphi(\frac{t}{n}) - 1} = (\log z)'\big|_{z=1} = 1.$$

(b) By (ii), $\varphi\left(\frac{t}{n}\right)^n \to e^{iat}$. By taking logarithm, we get

$$n\log\varphi\left(\frac{t}{n}\right) \to iat.$$

By (a), (b), we get

$$
\begin{aligned}
n\left(\varphi\left(\frac{t}{n}\right)-1\right) &= \frac{\varphi(\frac{t}{n})-1}{\log\varphi(\frac{t}{n})}\cdot n\log\varphi\left(\frac{t}{n}\right)\\
&\to iat.
\end{aligned}
$$

By Exercise 3.3.13, $\frac{S_n}{n}\Rightarrow a$ implies that $\varphi(\frac{t}{n})^n \to e^{iat}$ uniformly on a compact set. Fix small $0<\delta<1$. For any $h>0$, we can express $h=\frac{t}{n}$ with large n and $t\in[1-\delta,1+\delta]$. Then,

$$
\begin{aligned}
\left|\frac{\varphi(h)-1}{h}-ia\right| &= \left|\frac{\varphi(\frac{t}{n})-1}{t/n}-ia\right|\\
&= \left|\frac{1}{t}(n(\varphi(t/n)-1)-iat)\right|\\
&\le \sup_{t'\in[1-\delta,1+\delta]}\left|\frac{1}{t'}\right|\cdot|n(\varphi(t'/n)-1)-iat'|\\
&\le \frac{1}{1-\delta}\sup_{t'\in[1-\delta,1+\delta]}|n(\varphi(t'/n)-1)-iat'|.
\end{aligned}
$$

Since $n(\varphi(t/n)-1)\to iat$ uniformly on compact set,

$$
\limsup_{h\to 0}\left|\frac{\varphi(h)-1}{h}-ia\right| \le \frac{1}{1-\delta}\cancel{\limsup_{n\to\infty}\left[\sup_{t\in[1-\delta,1+\delta]}|n(\varphi(t/n)-1)-iat|\right]} = 0.
$$

Therefore, $\frac{\varphi(h)-1}{h}\to ia$ as $h\downarrow 0$.

Exercise 3.3.15. From Example 3.3.15, $\int_{-\infty}^{\infty}\frac{1-\cos x}{\pi x^2}dx = 1$.
Let $x=|y|t$. Then,

$$
\int_{-\infty}^{\infty}\frac{1-\cos|y|t}{\pi|y|^2t^2}|y|dt = \frac{1}{|y|}\int_{-\infty}^{\infty}\frac{1-\cos yt}{\pi t^2}dt = 1.
$$

Thus, $2\int_0^{\infty}\frac{1-\cos yt}{\pi t^2}dt = |y|$.
Now, integrate with repect to $F(y)$. Then, we get

$$
2\int_{-\infty}^{\infty}\int_0^{\infty}\frac{1-\cos yt}{\pi t^2}dtdF(y) = \int_{-\infty}^{\infty}|y|dF(y).
$$

Since $\frac{1-\cos yt}{\pi t^2}\ge 0$, by Fubini,

$$
2\int_0^{\infty}\int_{-\infty}^{\infty}\frac{1-\cos yt}{\pi t^2}dF(y)dt = 2\int_0^{\infty}\frac{1-Re(\varphi(t))}{\pi t^2}dt = \int_{-\infty}^{\infty}|y|dF(y).
$$

Exercise 3.3.16. Since $\lim_{t\downarrow 0}\frac{\varphi(t)-1}{t^2} = c > -\infty$,

$$
\lim_{t\downarrow 0}\overline{\left(\frac{\varphi(t)-1}{t^2}\right)} = \lim_{t\downarrow 0}\frac{\overline{\varphi(t)}-1}{t^2} = \lim_{t\downarrow 0}\frac{\varphi(-t)-1}{t^2} = c.
$$

Thus,

$$
\lim_{t\downarrow 0}\frac{\varphi(t)+\varphi(-t)-2\varphi(0)}{t^2} = 2c > -\infty.
$$

By Theorem 3.3.21, $EX^2 < \infty$.
By Theorem 3.3.20, $\varphi(t) = 1 + itEX - \frac{t^2}{2}EX^2 + o(t^2)$.
Then,

$$\begin{aligned}
\lim_{t\downarrow 0} \frac{\varphi(t)-1}{t^2} &= \lim_{t\downarrow 0} \frac{itEX - \frac{t^2}{2}EX^2 + o(t^2)}{t^2} \\
&= \lim_{t\downarrow 0} \frac{iEX}{t} - \frac{1}{2}EX^2 \\
&= c > -\infty.
\end{aligned}$$

Thus, $EX = 0$ and $EX^2 = -2c < \infty$.
Especially, if $\varphi(t) = 1 + o(t^2)$, we get $\lim_{t\downarrow 0} \frac{\varphi(t)-1}{t^2} = \lim_{t\downarrow 0} \frac{o(t^2)}{t^2} = 0$. Then, $EX = 0$ and $EX^2 = 0$. Thus, $\mathrm{Var}(X) = 0$, which means that X is constant a.s. Therefore, $X = 0$ a.s. Thus, $\varphi(t) = Ee^{itX} = 1$.

Exercise 3.3.17.

($\Rightarrow$) By Continuity Theorem, $\varphi_n(t) \to \varphi_0(t)$, for all t, where φ_0 is ch.f of 0. Since $\varphi_0(t) = Ee^{it0} = 1$, $\varphi_n(t) \to 1$, for all t.

($\Leftarrow$) **Step 1.** $\{Y_n\}$ is tight.

Let $f_n(t) = \frac{1}{\alpha}(1-\varphi_n(t))I_{[-\alpha,\alpha]}(t)$ for $\alpha < \delta$. Then, $\begin{cases} |f_n(t) \le \frac{2}{\alpha}I_{[-\alpha,\alpha]}(t) : \text{ integrable} \\ f_n(t) \to 0 \text{ by assumption} \end{cases}$.

By DCT, we have

$$\int f_n(t) = \frac{1}{\alpha}\int_{-\alpha}^{\alpha}(1-\varphi_n(t))dt \to 0.$$

Note that

$$\begin{aligned}
\frac{1}{u}\int_{-u}^{u}(1-\varphi_n(t))dt &= \frac{1}{u}\int_{-u}^{u}\left[1 - \int_{-\infty}^{\infty} e^{ity}d\mu_n(y)\right]dt \\
&= \frac{1}{u}\int_{-\infty}^{\infty}\int_{-u}^{u} 1 - e^{ity}dtd\mu_n(y) \\
&= \frac{1}{u}\int_{-\infty}^{\infty}\left[t - \frac{1}{iy}e^{ity}\right]_{-u}^{u} d\mu_n(y) \\
&= \frac{1}{u}\int_{-\infty}^{\infty}\left[\left(u - \frac{1}{iy}(\cos uy + i\sin uy)\right)\right. \\
&\qquad \left. - \left(-u - \frac{1}{iy}(\cos uy - i\sin uy)\right)\right]d\mu_n(y) \\
&= \frac{1}{u}\int_{-\infty}^{\infty}\left[2u - \frac{2}{y}\sin uy\right]d\mu_n(y) \\
&= 2\int_{-\infty}^{\infty}\left[1 - \frac{\sin uy}{uy}\right]d\mu_n(y) \\
&\ge 2\int_{|y|\ge\frac{2}{u}}\left[1 - \frac{1}{u|y|}\right]d\mu_n(y)
\end{aligned}$$

$$
\begin{aligned}
&\geq \int_{|y|\geq\frac{2}{u}} 1 d\mu_n(y) \\
&= \mu_n\left\{y : |y| \geq \frac{2}{u}\right\} \\
&\geq \mu_n\left\{y : |y| > \frac{2}{u}\right\}.
\end{aligned}
$$

The second equality holds by Fubini. The first inequality holds since $\frac{\sin uy}{uy} \leq 1$ and $\frac{\sin uy}{uy} \leq \frac{1}{u|y|}$.

Step 2. $Y_n \Rightarrow 0$.

Fix a subsequence $\{Y_{n_k}\}$. By Helly and Tightness Theorem, there exists sub-subsequence $\{Y_{n_{m(k)}}\}$ such that $Y_{n_{m(k)}} \Rightarrow X$. Since $\varphi_{n_{m(k)}} \to 1$ for $|t| < \delta$, $\varphi_X(t) = 1$ for $|t| < \delta$. Then, for $|t| < \delta$, $\frac{\varphi_X(t)-1}{t^2} = 0$, and thus

$$\lim_{t\downarrow 0} \frac{\varphi(t)-1}{t^2} = 0.$$

By Exercise 3.3.16, $\varphi_X(t) \equiv 1$. This means that $X \equiv 0$. Since this holds for all subsequence, by Theorem 3.2.15, $Y_n \Rightarrow 0$.

Exercise 3.3.18. Step 1. $S_m - S_n \xrightarrow{P} 0$.

Let $\varphi(t)$ be ch.f of X_1. Then, ch.f of S_n is given by $\varphi(t)^n$ and it converges to for some ch.f $\psi(t)$. Since ψ is ch.f, there exists $\delta > 0$ such that $\psi(t) > 0$ for $|t| \leq \delta$. Now, suppose $m > n$. ch.f of $S_m - S_n = \sum_{l=n+1}^{m} X_l$ is given by

$$\varphi(t)^{m-n} = \frac{\psi(t)^m}{\psi(t)^n}.$$

For $|t| \leq \delta$, as $m, n \to \infty$, $\varphi(t)^{m-n} \to 1$.
By Exercise 3.3.17, $S_m - S_n \Rightarrow 0$.
By Exercise 3.2.12, $S_m - S_n \xrightarrow{P} 0$.

Step 2. S_n converges in prob.
By Exercise 2.3.6,

$$d(S_n - S_m, 0) \to 0 \iff d(S_n, S_m) \to 0.$$

By Exercise 2.3.7, there exists a r.v. S_∞ such that $S_n \xrightarrow{P} S_\infty$.

Exercise 19-25 ($\star$)

Exercise 3.3.19.

Exercise 3.3.20.

Exercise 3.3.21.

Exercise 3.3.22.

Exercise 3.3.23.

Exercise 3.3.24.

Exercise 3.3.25.

3.4. **Central Limit Theorems.**

Exercise 3.4.1. $\Phi(-1)$.

Exercise 3.4.2. (a) Let $\sigma^2 = \mathrm{Var}(X_i) \in (0,\infty)$. By CLT, we have $\frac{S_n}{\sqrt{n}\sigma} \Rightarrow Z \sim N(0,1)$. Then, for any M,

$$\limsup_n P\left(\frac{S_n}{\sqrt{n}} > M\right) = \limsup_n P\left(\frac{S_n}{\sqrt{n}\sigma} > \frac{M}{\sigma}\right) = P(Z > M/\sigma) > 0.$$

Therefore,

$$P\left(\frac{S_n}{\sqrt{n}} > M \;\; i.o.\right) \geq \limsup_n P\left(\frac{S_n}{\sqrt{n}} > M\right) > 0.$$

Note that $\left\{\limsup_n \frac{S_n}{\sqrt{n}} > M\right\} \in \mathcal{T}$: tail σ-algebra.

By Kolmogorov's 0-1 law, $P\left(\limsup_n \frac{S_n}{\sqrt{n}} > M\right) = 1$, for all M.

Thus, $\limsup_n \frac{S_n}{\sqrt{n}} = \infty$ a.s.

(b) Assume that $\frac{S_n}{\sqrt{n}} \xrightarrow{P} U$ for some U.

Then, as $m \to \infty$,

$$\left|\frac{S_{m!}}{\sqrt{m!}} - \frac{S_{(m+1)!}}{\sqrt{(m+1)!}}\right| \xrightarrow{P} 0.$$

Since $S_{m!}$ and $S_{(m+1)!} - S_{m!}$ are independent, we get

$$P\left(1 < \frac{S_{m!}}{\sqrt{m!}} < 2,\ \frac{S_{(m+1)!} - S_{m!}}{\sqrt{(m+1)!}} < -3\right) = P\left(1 < \frac{S_{m!}}{\sqrt{m!}} < 2\right) P\left(\frac{S_{(m+1)!} - S_{m!}}{\sqrt{(m+1)!}} < -3\right)$$

$$\to P(1 < T < 2)P(T < -3) > 0, \tag{2a}$$

where $T \sim N(0,\sigma^2)$. For the second term, $\frac{1}{\sqrt{(m+1)!}} S_{m!} = \frac{1}{\sqrt{(m+1)!}} \cdot \frac{S_{m!}}{\sqrt{m!}} \xrightarrow{P} 0$.

Also,

$$\begin{aligned}
P\left(1 < \frac{S_{m!}}{\sqrt{m!}} < 2,\ \frac{S_{(m+1)!} - S_{m!}}{\sqrt{(m+1)!}} < -3\right) &= P\left(1 < \frac{S_{m!}}{\sqrt{m!}} < 2,\ \frac{S_{(m+1)!}}{\sqrt{(m+1)!}} < -3 + \frac{S_{m!}}{\sqrt{(m+1)!}}\right) \\
&\leq P\left(1 < \frac{S_{m!}}{\sqrt{m!}} < 2,\ \frac{S_{(m+1)!}}{\sqrt{(m+1)!}} < -3 + \frac{S_{m!}}{\sqrt{m!}}\right) \\
&\leq P\left(1 < \frac{S_{m!}}{\sqrt{m!}},\ \frac{S_{(m+1)!}}{\sqrt{(m+1)!}} < -1\right) \\
&\leq P\left(\left|\frac{S_{m!}}{\sqrt{m!}} - \frac{S_{(m+1)!}}{\sqrt{(m+1)!}}\right| > 2\right).
\end{aligned} \tag{2b}$$

By combining (2a) and (2b), we get

$$\liminf_m P\left(\left|\frac{S_{m!}}{\sqrt{m!}} - \frac{S_{(m+1)!}}{\sqrt{(m+1)!}}\right| > 2\right) > 0,$$

which contradicts to the assumption that $\frac{S_{m!}}{\sqrt{m!}} - \frac{S_{(m+1)!}}{\sqrt{(m+1)!}} \xrightarrow{P} 0$.

Exercise 3.4.3.

$$\begin{cases} X_i' : \text{independent copy of } X_i \\ Y_i = X_i - X_i' \\ U_i = Y_i \mathbf{1}_{(|Y_i| \le A)} \\ V_i = Y_i \mathbf{1}_{(|Y_i| > A)} \end{cases}$$

We know that $(X_i), (X_i'), (Y_i), (U_i), (V_i)$ are $i.i.d$ sequences. Suppose $EX_i^2 = \infty$.
Since there exists some r.v. W such that $\frac{S_n}{\sqrt{n}} \Rightarrow W$, $S_n' = \sum_{i=1}^n X_i'$ also weakly converges to W. That is, for all t, $\varphi_W(t) = \lim_{n\to\infty} \varphi_{S_n/\sqrt{n}}(t) = \lim_{n\to\infty} \varphi_{S_n'/\sqrt{n}}(t)$. Then, ch.f of $\frac{\sum_{m=1}^n Y_m}{\sqrt{n}}$ is given by

$$\begin{aligned} \varphi_{\frac{\sum_{m=1}^n Y_m}{\sqrt{n}}}(t) &= E \exp\left(it \left(\frac{S_n}{\sqrt{n}} - \frac{S_n'}{\sqrt{n}} \right) \right) \\ &= \varphi_{\frac{S_n}{\sqrt{n}}}(t) \cdot \varphi_{\frac{S_n'}{\sqrt{n}}}(-t) \\ &\to \varphi_W(t) \cdot \varphi_W(-t), \ \forall t. \end{aligned}$$

Since $\varphi_W(t) \cdot \varphi_W(-t)$ is continuous at 0, by Continuity Theorem, there exists a r.v. T such that $\frac{\sum_{m=1}^n Y_m}{\sqrt{n}} \Rightarrow T$. By hint, for any K, which is a continuity point of d.f of T,

$$P(Y \ge K) = \lim_{n\to\infty} P\left(\frac{\sum_{m=1}^n Y_m}{\sqrt{n}} \ge K \right) \ge \frac{1}{5}.$$

This means that $\lim_{K\to\infty} P(T \ge K) \ge \frac{1}{5}$ since discontinuity points of d.f of T is at most countable, which is a contradiction. One can derive hint from CLT and MCT.

Exercise 3.4.4. Let $S_n = X_1 + \cdots + X_n$. CLT for $i.i.d$ sequence implies that

$$\frac{S_n - ES_n}{\sqrt{n}\sigma} = \frac{S_n - n}{\sqrt{n}\sigma} \Rightarrow \chi. \tag{4a}$$

Note that $\frac{S_n - n}{\sqrt{n}\sigma} = \frac{\sqrt{S_n} + \sqrt{n}}{\sqrt{n}\sigma} \cdot (\sqrt{S_n} - \sqrt{n})$.
By WLLN, we have $\frac{S_n}{n} \xrightarrow{P} 1$. $(EX_1 = 1, \mathrm{Var}(X_1) < \infty)$.
By Continuous Mapping Theorem,

$$\frac{1}{\sigma} \left(\frac{\sqrt{S_n}}{\sqrt{n}} + 1 \right) \xrightarrow{P} \frac{2}{\sigma}. \tag{4b}$$

By Converging Together Lemma with (4a) and (4b),

$$\sqrt{S_n} - \sqrt{n} = \left(\frac{S_n - n}{\sqrt{n}\sigma} \right) \cdot \left(\frac{1}{\sigma} \left(\frac{\sqrt{S_n}}{\sqrt{n}} + 1 \right) \right)^{-1} \Rightarrow \frac{\sigma}{2}\chi$$

Exercise 3.4.5. By CLT,

$$\frac{\sum_{m=1}^n X_m}{\sqrt{n}\sigma} \Rightarrow \chi. \tag{5a}$$

By WLLN,

$$\frac{\sum_{m=1}^n X_m^2}{n} \xrightarrow{P} \sigma^2.$$

By Continuous Mapping Theorem,

$$\text{(5b)} \qquad \sqrt{\frac{\sum_{m=1}^n X_m^2}{n}} \xrightarrow{P} \sigma^2.$$

By Converging Together Lemma with (5a) and (5b),

$$\frac{\sum_{m=1}^n X_m}{\left(\sum_{m=1}^n X_m^2\right)^{1/2}} = \frac{\sum_{m=1}^n X_m}{\sqrt{n}\sigma} \cdot \sqrt{\frac{n\sigma^2}{\sum_{m=1}^n X_m^2}} \Rightarrow \chi.$$

Exercise 3.4.6. By CLT for *i.i.d* sequence, we know that $\frac{S_n}{\sigma\sqrt{n}} \Rightarrow N(0,1)$. Thus, its subsequence $\frac{S_{a_n}}{\sigma\sqrt{a_n}} \Rightarrow N(0,1)$. For Y_n and Z_n defined above, if one can show that $Y_n - Z_n \xrightarrow{P} 0$,

$$\frac{S_{N_n}}{\sigma\sqrt{a_n}} = \frac{S_{a_n}}{\sigma\sqrt{a_n}} + \frac{S_{a_n} - S_{N_n}}{\sigma\sqrt{a_n}} = \frac{S_{a_n}}{\sigma\sqrt{a_n}} + (Y_n - Z_n) \Rightarrow N(0,1)$$

by Converging Together Lemma. **Claim.** $Y_n - Z_n \xrightarrow{P} 0$.
Fix $\epsilon > 0$. Since $\frac{N_n}{a_n} \xrightarrow{P} 1$, $P\left(\left|\frac{N_n}{a_n} - 1\right| > \epsilon\right) = P\big(|N_n - a_n| > a_n\epsilon\big) \to 0$. Then,

$$\begin{aligned}
\limsup_n P(|Y_n - Z_n| > \delta) &= \limsup_n \Big[P(|Y_n - Z_n| > \delta,\ |N_n - a_n| > a_n\epsilon) + P(|Y_n - Z_n| > \delta,\ |N_n - a_n| \le a_n\epsilon)\Big] \\
&\le \limsup_n P(|Y_n - Z_n| > \delta,\ |N_n - a_n| > a_n\epsilon) + \limsup_n P(|Y_n - Z_n| > \delta,\ |N_n - a_n| \le a_n\epsilon) \\
&= \limsup_n P(|Y_n - Z_n| > \delta,\ |N_n - a_n| \le a_n\epsilon).
\end{aligned}$$

Note that

$$P(|Y_n - Z_n| > \delta, |N_n - a_n| \le a_n\epsilon) = P(|S_{N_n} - S_{a_n}| > \delta \cdot \sigma\sqrt{a_n},\ |N_n - a_n| \le a_n\epsilon).$$

Since $|N_n - a_n| \le a_n\epsilon$,

$$\begin{aligned}
P\Big[|S_{N_n} - S_{a_n}| > \delta\sigma\sqrt{a_n},\ |N_n - a_n| \le a_n\epsilon\Big] &\le P\left[\max_{(1-\epsilon)a_n \le m \le (1+\epsilon)a_n} |S_m - S_{[(1-\epsilon)a_n]}| > \frac{\delta\sigma\sqrt{a_n}}{2}\right] \\
&\le \frac{4}{\delta^2\sigma^2 a_n} \operatorname{Var}\left(S_{[(1+\epsilon)a_n]} - S_{[(1-\epsilon)a_n]}\right) \\
&= \frac{4}{\delta^2\sigma^2 a_n}\sigma^2\Big([(1+\epsilon)a_n] - [(1-\epsilon)a_n]\Big) \\
&\le \frac{8\epsilon}{\delta^2}
\end{aligned}$$

The second inequality holds by Kolmogorov's Maximal Inequality.
Thus, $\limsup_n P(|Y_n - Z_n > \delta) \le \frac{8\epsilon}{\delta^2}$. Since ϵ was arbitrarily chosen, $P(|Y_n - Z_n| > \delta) \to 0$, for any $\delta > 0$. Therefore, $Y_n \xrightarrow{P} Z_n$.

Exercise 3.4.7. 7

Exercise 3.4.8. 8

Exercise 3.4.9. (1) $\operatorname{Var}(S_n)/n \to 2$.

$\operatorname{Var}(X_n) = EX_m^2 = 2 - \frac{1}{m^2}$. Then, $\operatorname{Var}(S_n) = 2n - \sum_{m=1}^n \frac{1}{m^2}$. Since $\sum_{m=1}^n \frac{1}{m^2}$ converges, we get the desired result.

(2) $S_n/\sqrt{n} \Rightarrow \chi$.

Define $Y_m = \begin{cases} 1 & , X_m > 0 \\ -1 & , X_m < 0 \end{cases}$. Then, Y_m's are *i.i.d* sequence with $EY_m = 0$ and $\text{Var}(Y_m) = 1$. By CLT for *i.i.d*, we have

$$\sum_{m=1}^{n} Y_m \Big/ \sqrt{n} \Rightarrow \chi.$$

Note that $P(X_m \neq Y_m) = P(X_m = \pm m) = \frac{1}{m^2}$. Then, $\sum_{m=1}^{\infty} P(X_m \neq Y_m) = \sum_{m=1}^{\infty} \frac{1}{m^2} < \infty$. By Borel-Cantelli Lemma 1, $P(X_m \neq Y_m \ \ i.o.) = 0$. Thus, $P(X_m = Y_m \ \ eventually) = 1$. This means that

$$\frac{S_n}{\sqrt{n}} = \frac{\sum_{m=1}^{n} Y_m}{\sqrt{n}} + \frac{S_n - \sum_{m=1}^{n} Y_m}{\sqrt{n}} \Rightarrow \chi$$

by converging together Lemma since the second term converges to 0.

Exercise 3.4.10. Let $s_n^2 = \sum_{m=1}^{n} \text{Var}(X_m)$ and $Y_{nm} = \frac{X_m - EX_m}{s_n}$. Then, $EY_{mn} = 0$ and $t_n^2 := \sum_{m=1}^{n} \text{Var}(Y_{nm}) = 1$. For Y_{nm},

(i) $\sum_{m=1}^{n} EY_{nm}^2 = t_n^2 = 1$

(ii) For $\epsilon > 0$, $|Y_{nm}| > \epsilon \iff |X_m - EX_m| > \epsilon s_n$.
Note that since $|X_i| \leq M$, $|X_m - EX_m| \leq 2M$, $\forall m$. However, since $\sum_{n=1}^{\infty} \text{Var}(X_n) = \infty$, $\epsilon S_n > 2M$ for sufficiently large n. This means that $\{|Y_{nm}| > \epsilon\}$ does not occur eventually. Thus,

$$\lim_{n\to\infty} \sum_{m=1}^{n} \int_{|Y_{nm}|>\epsilon} |Y_{nm}|^2 dP = 0.$$

By (i) and (ii), Lindeberg Theorem gives us that $\sum_{m=1}^{n} Y_{nm} \Rightarrow \chi$. Therefore,

$$\frac{\sum_{m=1}^{n} X_m - E\left(\sum_{m=1}^{n} X_m\right)}{\sqrt{\sum_{m=1}^{n} \text{Var}(X_m)}} = \frac{S_n - ES_n}{\sqrt{\text{Var}(S_n)}} \Rightarrow \chi.$$

Exercise 3.4.11. By Exercise 3.4.12, it suffices to show that $\{X_n\}$ satisfies Lyapounov's condition.

$$\begin{aligned}
&\lim_{n\to\infty} \alpha_n^{-(2+\delta)} \sum_{m=1}^{n} E(|X_m - EX_m|^{2+\delta}) \\
=& \lim_{n\to\infty} \frac{1}{n^{1+\delta/2}} \sum_{m=1}^{n} E|X_m|^{2+\delta} \\
=& \lim_{n\to\infty} \frac{C}{n^{\delta/2}} \\
=& 0.
\end{aligned}$$

Exercise 3.4.12. Let $X_{n,m} = \frac{X_m - EX_m}{\alpha_n}$. It suffices to show that $\{X_{n,m}\}$ satisfies the Lindeberg condition.

(i) $\sum_{m=1}^{n} EX_{n,m}^2 = 1$.

(ii) For given $\epsilon > 0$,

$$\begin{aligned}\sum_{m=1}^{n} E(X_{n,m}^2; |X_{n,m}| > \epsilon) &= \sum_{m=1}^{n} \int_{|X_{n,m}|>\epsilon} X_{n,m}^2 dP \\ &= \frac{1}{\alpha_n^2} \sum_{m=1}^{n} \int_{|X_m - EX_m| > \epsilon\alpha_n} |X_m - EX_m|^2 dP \\ &\le \frac{1}{\alpha_n^2} \sum_{m=1}^{n} \int_{|X_m - EX_m| > \epsilon\alpha_n} \left(\frac{|X_m - EX_m|}{\epsilon\alpha_n} \right)^{\delta} |X_m - EX_m|^2 dP \\ &\le \frac{1}{\epsilon^\delta} \cdot \left[\frac{1}{\alpha_n^{2+\delta}} \sum_{m=1}^{n} \int |X_m - EX_m|^{2+\delta} dP \right] \to 0\end{aligned}$$

by assumption.

By (i) and (ii), we get the desired result.

Exercise 3.4.13. (i) $\beta > 1$.

$P(X_j \neq 0) = P(X_j = \pm j) = \frac{1}{j^\beta}$. Thus,

$$\sum_{j=1}^{\infty} P(X_j \neq 0) = \sum_{j=1}^{\infty} \frac{1}{j^\beta} < \infty.$$

By Borel-Cantelli Lemma 1, $P(X_j = 0 eventually) = 1$. For $\omega \in \Omega_0 := \{\omega : X_j(\omega) = 0 \ eventually\}$, $\sum_{j=1}^{\infty} X_j(\omega)$ converges. Thus, $S_n \to S_\infty$ a.s. , where S_∞ is a r.v. such that $S_\infty(\omega) = \sum_{j=1}^{\infty} X_j(\omega)$ for $\omega \in \Omega_0$

(ii) $\beta < 1$.

Let

$$\begin{cases} X_{nm} &= X_m \text{ for } 1 \le m \le n, \ \forall n \\ Y_{nm} &= \frac{X_{nm}}{n^{(3-\beta)/2}} \\ T_n &= \sum_{m=1}^{n} Y_{nm} = \frac{S_n}{n^{(3-\beta)/2}} \end{cases}$$

Then, $EY_{nm} = 0$ and $EY_{nm}^2 = \frac{1}{n^{(3-\beta}} m^{2-\beta}$.

(a) $\sum_{m=1}^{n} Y_{nm}^2 = \frac{1}{n^{(3-\beta)}} \sum_{m=1}^{n} m^{2-\beta} \le \frac{1}{n^{(3-\beta)}} \cdot n \cdot n^{2-\beta} = 1$.
Since $\sum_{m=1}^{n} Y_{nm}^2$ is bounded and monotonically increasing, $\sum_{m=1}^{n} Y_{nm}^2 \to c^2 \le 1$ for some $c > 0$.

(b) $\lim_{n\to\infty} \sum_{m=1}^{n} E(|Y_{nm}|^2; |Y_{nm}| > \epsilon) = \lim_{n\to\infty} \sum_{m=1}^{n} E(|Y_{nm}|^2; |X_{nm}| > \epsilon \cdot n^{(3-\beta)/2})$.
Since $\beta < 1$, $\frac{3-\beta}{2} > 1$. Then, for large n, $\epsilon \cdot n^{(3-\beta)/2} > n$.Then, $\left\{|X_{nm}| > \epsilon \cdot n^{(3-\beta)/2}\right\} = \emptyset$. Thus, $\lim_{n\to\infty} \sum_{m=1}^{n} E(|Y_{nm}|^2; |X_{nm}| > \epsilon \cdot n^{(3-\beta)/2}) = 0$.

By (a) and (b), Lindeberg Theorem implies that $T_n = \frac{S_n}{n^{(3-\beta)/2}} \Rightarrow c\chi$.

(iii) $\beta = 1$.

$$\varphi_{S_n/n}(t) = \prod_{j=1}^{n} \varphi_{X_j}(t/n)$$

$$= \prod_{j=1}^{n} \left(1 - \frac{1}{j}\left(1 - \cos\frac{jt}{n}\right)\right)$$
$$= \prod_{j=1}^{n} \left(1 - \frac{1}{n} \cdot \left(\frac{j}{n}\right)^{-1} \cdot \left(1 - \cos\frac{jt}{n}\right)\right).$$

Note that
$$\sum_{j=1}^{n} \frac{1}{n}\left(\frac{j}{n}\right)^{-1}\left(1 - \cos\frac{jt}{n}\right) \to \int_0^1 \frac{1}{x}(1 - \cos xt)dx.$$

Also, $\frac{1}{j}\left(\cos\frac{jt}{n} - 1\right) \le 0$ for all j and $\lim_{n\to\infty} \frac{1}{j}\left(\cos\frac{jt}{n} - 1\right) \to 0$.
Thus, $\max_{1\le j\le n} \frac{1}{j}\left(\cos\frac{jt}{n} - 1\right) \to 0$.
Exercise **??** with $c_{j,n} = -\frac{1}{j}\left(1 - \cos\frac{jt}{n}\right)$ implies

$$\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \prod_{j=1}^{n}\left(1 + \left(-\frac{1}{j}\left(1 - \cos\frac{jt}{n}\right)\right)\right)$$
$$\to \exp\left[-\int_0^1 \frac{1}{x}(1 - \cos tx)dx\right].$$

Since the last one is the ch.f of χ, it is continuous at $t = 0$. By Continuity theorem, we get $S_n/n \Rightarrow \chi$.

3.5. **Local Limit Theorems.** No exercise

3.6. **Poisson Convergence.** No exercise

3.7. **Poisson Processes.**

Exercise 3.7.1. 1

Exercise 3.7.2. 2

Exercise 3.7.3. 3

Exercise 3.7.4. 4

Exercise 3.7.5. 5

Exercise 3.7.6. 6

Exercise 3.7.7. 7

Exercise 3.7.8. 8

Exercise 3.7.9. 9

3.8. **Stable Laws*.**

Exercise 3.8.1. 1

Exercise 3.8.2. 2

Exercise 3.8.3. 3

Exercise 3.8.4. 4

Exercise 3.8.5. 5

Exercise 3.8.6. 6

Exercise 3.8.7. 7

3.9. **Infinitely Divisible Distributions*.**

Exercise 3.9.1. 1

Exercise 3.9.2. 2

Exercise 3.9.3. 3

Exercise 3.9.4. 4

3.10. **Limit Theorems in $\mathbb{R}^d$.**

Exercise 3.10.1.

$$\begin{aligned}
P(X_i \le x) &= P(X_i \le x, X_j \in (-\infty, \infty) \text{ for } j \ne i) \\
&= \lim_{y\to\infty} P(X_1 \le y, \cdots, X_{i-1} \le y, X_i \le x, X_{i+1} \le y, \cdots, X_d \le y) \\
&= \lim_{y\to\infty} F(y, \cdots, y, x, y, \cdots, y),
\end{aligned}$$

where x lies at i-th coordinate.

Exercise 3.10.2. Omit.

Exercise 3.10.3. Omit.

Exercise 3.10.4. Portmanteu

Exercise 3.10.5. Let X be a r.v. whose ch.f is φ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a random vector whose ch.f is ψ.
Then,

$$\begin{aligned}
\varphi(t_1 + \cdots + t_d) &= E\exp[i(t_1 + \cdots t_d)X] \\
\psi(t_1, \ldots, t_d) &= E\exp[i\mathbf{t}^T\mathbf{Y}] \\
&= E\exp[i(t_1 Y_1 + \cdots + t_d Y_d)].
\end{aligned}$$

Thus, $\mathbf{Y} = (X, \ldots, X)$

Exercise 3.10.6. Let $\mathbf{X} = (X_1, \ldots, X_k)^T$.

($\Rightarrow$)
$$\begin{aligned}
\varphi_{X_1,\cdots,X_k}(t) &= Ee^{i\mathbf{t}^T\mathbf{X}} \\
&= Ee^{it_1X_1 + \cdots + t_kX_k} \\
&= Ee^{it_1X_1} \cdots Ee^{it_kX_k} \\
&= \prod_{j=1}^k \varphi_{X_j}(t_j)
\end{aligned}$$

The third equality holds since X_i's are independent.

($\Leftarrow$) For $A = [a_1, b_1] \times \cdots \times [a_k, b_k]$ with $\mu_X(\partial A) = 0$,

$$\begin{aligned}
\mu_{\mathbf{X}}(A) &= \lim_{T\to\infty}\left[(2\pi)^{-k}\int_{[-T,T]^k}\left(\prod_{j=1}^k \psi_j(t_j)\right)\varphi_{\mathbf{X}}(\mathbf{t})d\mathbf{t}\right] \\
&= \lim_{T\to\infty}\left[(2\pi)^{-k}\int_{[-T,T]^k}\prod_{j=1}^k\Big(\psi_j(t_j)\varphi_{X_j}(t_j)\Big)dt_1\cdots dt_j\right] \\
&= \lim_{T\to\infty}\left[\prod_{j=1}^k\left((2\pi)^{-1}\int_{[-T,T]}\psi_j(t_j)\varphi_{X_j}(t_j)dt_j\right)\right]
\end{aligned}$$

$$= \prod_{j=1}^{k} \left[\lim_{T\to\infty} (2\pi)^{-1} \int_{[-T,T]} \psi_j(t_j)\varphi_{X_j}(t_j)dt_j \right]$$
$$= \prod_{j=1}^{k} \mu_{X_j}\big([a_j, b_j]\big).$$

Since it holds for any A, $X_1, \cdots, X_k$ are independent.

Exercise 3.10.7. Let $\mathbf{X} = (X_1, \ldots, X_d)^T$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)^T$ and $\boldsymbol{\Gamma} = (\Gamma_{ij})$

($\Rightarrow$)
$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left[i\mathbf{t}^T\boldsymbol{\theta} - \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Gamma_{ij}t_it_j \right]$$
$$= E\exp(i\mathbf{t}^T\mathbf{X})$$
$$= E\exp(it_1X_1 + \cdots it_dX_d)$$
$$= E\exp(it_1X_1)\cdots E\exp(it_dX_d)$$
$$= \exp\left[i\mathbf{t}^T\boldsymbol{\theta} - \frac{1}{2}\sum_{l=1}^{d}\Gamma_{ll}t_l^2 \right]$$

The fourth equality holds since X_i's are independent. The fifth equality comes from Exercise 3.10.8 with $\mathbf{c} = \mathbf{e}_i$. Thus, $\Gamma_{ij} = 0$ for $i \neq j$.

($\Leftarrow$) Similarly,

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left[i\mathbf{t}^T\boldsymbol{\theta} - \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Gamma_{ij}t_it_j \right]$$
$$= \exp\left[i\mathbf{t}^T\boldsymbol{\theta} - \frac{1}{2}\sum_{l=1}^{d}\Gamma_{ll}t_l^2 \right]$$
$$= \prod_{i=1}^{d}\exp\left(it_i\theta_i - \frac{1}{2}\Gamma_{ii}t_i^2 \right)$$
$$= \prod_{i=1}^{d}\varphi_{X_i}(t_i)$$

The last equality comes from Exercise 3.10.8 with $\mathbf{c} = \mathbf{e}_i$.

Exercise 3.10.8. Let $\mathbf{X} = (X_1, \ldots, X_d)^T$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)^T, \mathbf{c} = (c_1, \ldots, c_d)^T$ and $\boldsymbol{\Gamma} = (\Gamma_{ij})$: $d \times d$ matrix.

We want to show that

$$X \sim N_d(\boldsymbol{\theta}, \boldsymbol{\Gamma}) \iff \forall \mathbf{c} \in \mathbb{R}^d, \mathbf{c}^T\mathbf{X} \sim N(\mathbf{c}^T\boldsymbol{\theta}, \mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}).$$

Note that ch.f of $\mathbf{X} = N_d(\boldsymbol{\theta}, \boldsymbol{\Gamma})$ is given by

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left[i\mathbf{t}^T\boldsymbol{\theta} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Gamma}\mathbf{t} \right].$$

$$
\begin{aligned}
(\Rightarrow)\ \varphi_{\mathbf{c}^T\mathbf{X}}(t) &= Ee^{it\mathbf{c}^T\mathbf{X}} \\
&= Ee^{i(t\mathbf{c})^T\mathbf{X}} \\
&= \exp\left[i(t\mathbf{c})^T\boldsymbol{\theta} - \frac{1}{2}(t\mathbf{c})^T\boldsymbol{\Gamma}(t\mathbf{c})\right] \\
&= \exp\left[it(\mathbf{c}^T\theta) - \frac{1}{2}(\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c})t^2\right],
\end{aligned}
$$

which is ch.f of $N(\mathbf{c}^T\boldsymbol{\theta}, \mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c})$. Thus, $\mathbf{c}^T\mathbf{X} = c_1X_1 + \cdots + c_dX_d \sim N(\mathbf{c}^T\boldsymbol{\theta}, \mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c})$.

$$
\begin{aligned}
(\Leftarrow)\ \varphi_{\mathbf{X}}(\mathbf{t}) &= Ee^{i\mathbf{t}^T\mathbf{X}} \\
&= Ee^{i\cdot 1\cdot(\mathbf{t}^T\mathbf{X})} \\
&= \varphi_{\mathbf{t}^T\mathbf{X}}(1) \\
&= \exp\left[i\cdot 1\cdot(\mathbf{t}^T\boldsymbol{\theta}) - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Gamma}\mathbf{t}\cdot 1^2\right] \\
&= \exp\left[i\mathbf{t}^T\boldsymbol{\theta} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Gamma}\mathbf{t}\right],
\end{aligned}
$$

which is ch.f of $N_d(\boldsymbol{\theta}, \boldsymbol{\Gamma})$. Thus, $\mathbf{X} \sim N_d(\boldsymbol{\theta}, \boldsymbol{\Gamma})$.

4. Martingales

4.1. **Conditional Expectation.**

Exercise 4.1.1. Since $G, \Omega \in \mathcal{G}$,

$$
\begin{aligned}
\int_G \mathbb{P}(A|\mathcal{G})\,d\mathbb{P} &= \int_G \mathbf{1}_A\,d\mathbb{P} = \mathbb{P}(A\cap G) \\
\int_\Omega \mathbb{P}(A|\mathcal{G})\,d\mathbb{P} &= \int_\Omega \mathbf{1}_A\,d\mathbb{P} = \mathbb{P}(A)
\end{aligned}
$$

and the result follows.

Exercise 4.1.2. Let $F \in \mathcal{F}$.

$$
\int_F \mathbb{P}(|X|\geq a|\mathcal{F})\,d\mathbb{P} = \int_F \mathbf{1}_{|X|\geq a}\,d\mathbb{P} \leq \int_F \frac{|X|^2}{a^2}\,d\mathbb{P} = \int_F a^{-2}\mathbb{E}\left[|X|^2|\mathcal{F}\right]\,d\mathbb{P}.
$$

Since $F \in \mathcal{F}$ was arbitrary, the result follows.

Exercise 4.1.3. Almost surely, we have

$$
0 \leq \mathbb{E}^{\mathcal{G}}[(|X| + \theta|Y|)^2] = \mathbb{E}^{\mathcal{G}}[X^2] + 2\theta\mathbb{E}^{\mathcal{G}}|XY| + \theta^2\mathbb{E}^{\mathcal{G}}[Y^2].
$$

Thus, almost surely,

$$
D/4 = (\mathbb{E}^{\mathcal{G}}|XY|)^2 - \mathbb{E}^{\mathcal{G}}[X^2]\mathbb{E}^{\mathcal{G}}[Y^2] \leq 0.
$$

Exercise 4.1.4. Let $\mu : \Omega\times\mathcal{F} \to [0,1]$ be an r.c.p. given $\mathcal{G}$.

$$
\mathbb{E}^{\mathcal{G}}|XY| = \int |X(\omega)Y(\omega)|\mu(\tilde{\omega}, d\omega)
$$

$$
\begin{aligned}
&\leq \left(\int |X(\omega)|^p \mu(\tilde{\omega}, d\omega)\right)^{1/p} \left(\int |Y(\omega)|^q \mu(\tilde{\omega}, d\omega)\right)^{1/q} \\
&= (\mathbb{E}^{\mathcal{G}}|X|^p)^{1/p} (\mathbb{E}^{\mathcal{G}}|Y|^q)^{1/q}.
\end{aligned}
$$

Exercise 4.1.5. Let $\mathbb{P}(a) = \mathbb{P}(b) = \mathbb{P}(c) = 1/3$ and let $\mathcal{F}_1 = \sigma(\{a\})$, $\mathcal{F}_2 = \sigma(\{b\})$. Define X as

$$
X = \begin{cases} a \mapsto 1 \\ b \mapsto 2 \\ c \mapsto 3 \end{cases}.
$$

Then,

$$
\mathbb{E}^{\mathcal{F}_1} X = \begin{cases} a \mapsto 1 \\ b \mapsto 2.5 \\ c \mapsto 2.5 \end{cases}, \quad \mathbb{E}^{\mathcal{F}_2} X = \begin{cases} a \mapsto 2 \\ b \mapsto 2 \\ c \mapsto 2 \end{cases}
$$

and

$$
\mathbb{E}^{\mathcal{F}_2}\mathbb{E}^{\mathcal{F}_1} X = \begin{cases} a \mapsto 1.75 \\ b \mapsto 2.5 \\ c \mapsto 1.75 \end{cases}, \quad \mathbb{E}^{\mathcal{F}_1}\mathbb{E}^{\mathcal{F}_2} X = \begin{cases} a \mapsto 2 \\ b \mapsto 2 \\ c \mapsto 2 \end{cases}.
$$

Exercise 4.1.6.

$$
\begin{aligned}
&\mathbb{E}[(X - \mathbb{E}^{\mathcal{F}}X)^2] + \mathbb{E}[(\mathbb{E}^{\mathcal{F}}X - \mathbb{E}^{\mathcal{G}}X)^2] \\
&\quad = \mathbb{E}X^2 + \mathbb{E}[(\mathbb{E}^{\mathcal{F}}X)^2] - 2\mathbb{E}[X\mathbb{E}^{\mathcal{F}}X] + \mathbb{E}[(\mathbb{E}^{\mathcal{F}}X)^2] + \mathbb{E}[(\mathbb{E}^{\mathcal{G}}X)^2] - 2\mathbb{E}[(\mathbb{E}^{\mathcal{F}}X)(\mathbb{E}^{\mathcal{G}}X)].
\end{aligned}
$$

Note that $\mathbb{E}[(\mathbb{E}^{\mathcal{F}}X)^2] = \mathbb{E}[(\mathbb{E}^{\mathcal{F}}X)(\mathbb{E}^{\mathcal{F}}X)] = \mathbb{E}[\mathbb{E}^{\mathcal{F}}[X\mathbb{E}^{\mathcal{F}}X]] = \mathbb{E}[X\mathbb{E}^{\mathcal{F}}X]$, and $\mathbb{E}[(\mathbb{E}^{\mathcal{F}}X)(\mathbb{E}^{\mathcal{G}}X)] = \mathbb{E}[\mathbb{E}^{\mathcal{F}}[X\mathbb{E}^{\mathcal{G}}X]] = \mathbb{E}[X\mathbb{E}^{\mathcal{G}}X] = \mathbb{E}[(\mathbb{E}^{\mathcal{G}}X)^2]$. Thus the above becomes

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

Exercise 4.1.7. Note that

$$
\mathrm{Var}[X|\mathcal{F}] = \mathbb{E}^{\mathcal{F}}X^2 - (\mathbb{E}^{\mathcal{F}}X)^2 = \mathbb{E}^{\mathcal{F}}[(X - \mathbb{E}^{\mathcal{F}}X)^2].
$$

Thus,

$$
\mathbb{E}[\mathrm{Var}[X|\mathcal{F}]] = \mathbb{E}[(X - \mathbb{E}^{\mathcal{F}}X)^2].
$$

Also,

$$
\mathrm{Var}[\mathbb{E}^{\mathcal{F}}X] = \mathbb{E}[(\mathbb{E}^{\mathcal{F}}X - \mathbb{E}X)^2].
$$

Thus the result follows from Exercise 4.1.6.

Exercise 4.1.8. Let $\mathcal{F} = \sigma(N)$. First, we need to check that $\mathbb{E}|X|^2 < \infty$. Indeed,

$$\begin{aligned}
\mathbb{E}|X|^2 &= \sum_n \mathbb{E}[|X|^2 \mathbf{1}_{\{N=n\}}] \\
&= \sum_n \mathbb{E}[(Y_1 + \cdots + Y_n)^2 \mathbf{1}_{\{N=n\}}] \\
&= \sum_n \mathbb{E}[(Y_1 + \cdots + Y_n)^2]\mathbb{E}[\mathbf{1}_{\{N=n\}}] \\
&= \sum_n (n\sigma^2 + n^2\mu^2)\mathbb{P}[N = n] \\
&\le \sum_n Cn^2 \mathbb{P}[N = n] \\
&= C\mathbb{E}N^2 < \infty
\end{aligned}$$

where for the third equality we used the independence of Y_i and N. Now we claim that $\mathbb{E}^{\mathcal{F}} X = N\mu$ and $\mathrm{Var}[X|\mathcal{F}] = N\sigma^2$. Note that these are $\mathcal{F}$-measurable. For $A_n = \{N = n\}$, we have

$$\begin{aligned}
\int_{A_n} X \, d\mathbb{P} &= \int_{A_n} (Y_1 + \cdots + Y_n) \, d\mathbb{P} \\
&= \mathbb{E}[(Y_1 + \cdots + Y_n)\mathbf{1}_{A_n}] \\
&= \mathbb{E}[Y_1 + \cdots + Y_n]\mathbb{E}[\mathbf{1}_{A_n}] \\
&= n\mu\mathbb{E}[\mathbf{1}_{A_n}] \\
&= \int_{A_n} N\mu \, d\mathbb{P}
\end{aligned}$$

and

$$\begin{aligned}
\int_{A_n} X^2 \, d\mathbb{P} &= \int_{A_n} (Y_1 + \cdots + Y_n)^2 \, d\mathbb{P} \\
&= \mathbb{E}[(Y_1 + \cdots + Y_n)^2 \mathbf{1}_{A_n}] \\
&= \mathbb{E}[(Y_1 + \cdots + Y_n)^2]\mathbb{E}[\mathbf{1}_{A_n}] \\
&= (n\sigma^2 + n^2\mu^2)\mathbb{E}[\mathbf{1}_{A_n}] \\
&= \int_{A_n} (N\sigma^2 + N^2\mu^2) \, d\mathbb{P}.
\end{aligned}$$

Since $\mathcal{F}$ is generated by $A_n, n \ge 1$ by complement and countable disjoint union (cf. Dynkin's π-λ theorem), we have the above equalities for all $A \in \mathcal{F}$. Thus $\mathbb{E}^{\mathcal{F}} X = N\mu$ and $\mathrm{Var}[X|\mathcal{F}] = \mathbb{E}^{\mathcal{F}} X^2 - (\mathbb{E}^{\mathcal{F}} X)^2 = N\sigma^2$. Finally, from Exercise 4.1.7,

$$\mathrm{Var}[X] = \mathrm{Var}[\mathbb{E}^{\mathcal{F}} X] + \mathbb{E}[\mathrm{Var}[X|\mathcal{F}]] = \mu^2 \,\mathrm{Var}[N] + \sigma^2 \mathbb{E}N.$$

Exercise 4.1.9. From Exercise 4.1.7,

$$\mathbb{E}[(Y - X)^2] + \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[(Y - \mathbb{E}Y)^2].$$

Note that $\mathbb{E}[(X-\mathbb{E}X)^2] = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = \mathbb{E}[(Y-\mathbb{E}Y)^2]$, where the second equality holds since $\mathbb{E}X^2 = \mathbb{E}Y^2$ and $\mathbb{E}X = \mathbb{E}Y$. Thus $\mathbb{E}[(Y-X)^2] = 0$ and $X = Y$ a.s.

Exercise 4.1.10. Fix $c \in \mathbb{Q}$. Since $\mathbb{E}[Y|\mathcal{G}] - c \overset{d}{=} Y - c$, we have $\mathbb{E}|\mathbb{E}[Y|\mathcal{G}] - c| = \mathbb{E}|Y - c|$. Meanwhile, we also have

$$\mathbb{E}|\mathbb{E}[Y|\mathcal{G}] - c| = \mathbb{E}|\mathbb{E}[Y-c|\mathcal{G}]| \leq \mathbb{E}\mathbb{E}[|Y-c||\mathcal{G}] = \mathbb{E}|Y-c|$$

and thus

$$\mathbb{E}|\mathbb{E}[Y-c|\mathcal{G}]| = \mathbb{E}\mathbb{E}[|Y-c||\mathcal{G}] \quad \text{a.s.}$$

Since $|\mathbb{E}[Y-c|\mathcal{G}]| \leq \mathbb{E}[|Y-c||\mathcal{G}]$, it follows that

$$|\mathbb{E}[Y-c|\mathcal{G}]| = \mathbb{E}[|Y-c||\mathcal{G}] \quad \text{a.s.}$$

Now let $A = \{\mathbb{E}[Y-c|\mathcal{G}] > 0\} \in \mathcal{G}$. Multiplying $\mathbf{1}_A$ both sides, the last display becomes

$$\begin{aligned}
\mathbb{E}[(Y-c)\mathbf{1}_A|\mathcal{G}] &= \mathbb{E}[Y-c|\mathcal{G}]\mathbf{1}_A \\
&= |\mathbb{E}[Y-c|\mathcal{G}]|\mathbf{1}_A \\
&= \mathbb{E}[|Y-c||\mathcal{G}]\mathbf{1}_A \\
&= \mathbb{E}[|Y-c|\mathbf{1}_A|\mathcal{G}].
\end{aligned}$$

Since $(Y-c)\mathbf{1}_A \leq |Y-c|\mathbf{1}_A$, it again implies that $(Y-c)\mathbf{1}_A = |Y-c|\mathbf{1}_A$ a.s. A similar argument with $B = \{\mathbb{E}[Y-c|\mathcal{G}] \leq 0\}$ in place of A will show that $-(Y-c)\mathbf{1}_B = |Y-c|\mathbf{1}_B$ a.s. Thus $\operatorname{sgn}(Y-c) = \operatorname{sgn}(\mathbb{E}[Y-c|\mathcal{G}]) = \operatorname{sgn}(\mathbb{E}[Y|\mathcal{G}]-c)$. Now suppose that $\{Y \neq \mathbb{E}[Y|\mathcal{G}]\}$ has positive probability. Without loss of generality we may assume that $\{Y < \mathbb{E}[Y|\mathcal{G}]\}$ has a positive probability. Since $\{Y < \mathbb{E}[Y|\mathcal{G}]\} = \bigcup_{c\in\mathbb{Q}}\{Y < c < \mathbb{E}[Y|\mathcal{G}]\}$, there must exist some $c \in \mathbb{Q}$ such that $\{Y < c < \mathbb{E}[Y|\mathcal{G}]\}$ has a positive probability. This is a contradiction because then with a positive probability, $\operatorname{sgn}(Y-c) \neq \operatorname{sgn}(\mathbb{E}[Y|\mathcal{G}]-c)$. Therefore $Y = \mathbb{E}[Y|\mathcal{G}]$ a.s.

4.2. **Martingales, Almost Sure Convergence.**

Exercise 4.2.1. Since $X_1, \cdots, X_n$ are measurable with respect to $\mathcal{G}_n$, $\mathcal{F}_n \subset \mathcal{G}_n$. Note that $\mathbb{E}^{\mathcal{F}_n}[X_{n+1}] = \mathbb{E}^{\mathcal{F}_n}[\mathbb{E}^{\mathcal{G}_n}X_{n+1}] = \mathbb{E}^{\mathcal{F}_n}X_n = X_n$. Thus X_n is a martingale with respect to $\mathcal{F}_n$.

Exercise 4.2.2. Let $X_n = -1/n$. Then $\mathbb{E}^{\mathcal{F}_n}X_{n+1} = -1/(n+1) > -1/n = X_n$ and $\mathbb{E}^{\mathcal{F}_n}X_{n+1}^2 = 1/(n+1)^2 < 1/n^2 = X_n^2$.

Exercise 4.2.3. Note that, almost surely,

$$\begin{aligned}
\mathbb{E}^{\mathcal{F}_n}[X_{n+1} \vee Y_{n+1}] &= \mathbb{E}^{\mathcal{F}_n}[X_{n+1}\mathbf{1}_{\{X_{n+1}\geq Y_{n+1}\}}] + \mathbb{E}^{\mathcal{F}_n}[Y_{n+1}\mathbf{1}_{\{X_{n+1}<Y_{n+1}\}}] \\
&\geq \mathbb{E}^{\mathcal{F}_n}[Y_{n+1}\mathbf{1}_{\{X_{n+1}\geq Y_{n+1}\}}] + \mathbb{E}^{\mathcal{F}_n}[Y_{n+1}\mathbf{1}_{\{X_{n+1}<Y_{n+1}\}}] \\
&= \mathbb{E}^{\mathcal{F}_n}[Y_{n+1}] = Y_n,
\end{aligned}$$

and similarly $\mathbb{E}^{\mathcal{F}_n}[X_{n+1} \vee Y_{n+1}] \geq X_n$ a.s. Thus $\mathbb{E}^{\mathcal{F}_n}[X_{n+1} \vee Y_{n+1}] \geq X_n \vee Y_n$.

Exercise 4.2.4. Let $T_M = \inf\{n \le 0 \colon X_n^+ > M\}$. Consider the stopped process $X_{T_M \wedge n}$. Note that $X_{T_M\wedge n}^+ \le M + \sup_j \xi_j^+$, which leads to $\mathbb{E}X_{T_M\wedge n}^+ \le M + \mathbb{E}\sup_j \xi_j^+ < \infty$ for all n. Thus $X_{T_M\wedge n}$ converges a.s. That is, X_n converges on the event $\{T_M = \infty\}$. Since we have $\sup_n X_n < \infty$, if we let $M \to \infty$, $\{T_M = \infty\} \uparrow \Omega$ up to a measure zero set. Therefore, X_n converges a.s.

Exercise 4.2.5. Let ξ_j be defined as

$$\xi_j = \begin{cases} -1 & \text{with probability } \dfrac{1}{\epsilon_j + 1} \\ \dfrac{1}{\epsilon_j} & \text{with probability } \dfrac{\epsilon_j}{\epsilon_j + 1} \end{cases}$$

where $\{\epsilon_j\}$ satisfies $\sum_j \epsilon_j < \infty$. By the Borel-Cantelli lemma, $\mathbb{P}(\xi_j \ne -1 \text{ i.o.}) = 0$ since $\sum \mathbb{P}(\xi_j \ne -1) < \infty$. That is, $\xi_j = -1$ eventually a.s. and thus $X_n \to -\infty$ a.s.

Exercise 4.2.6. (i) Note that $X_n \ge 0$. Since X is a martingale, it is a supermartingale, and thus X_n converges a.s. to an L^1 random variable X. To see $X = 0$ a.s., note that

$$\begin{aligned} \mathbb{P}(|X_{n+1} - X_n| > \delta\epsilon) &= \mathbb{P}(X_n|Y_{n+1} - 1| > \delta\epsilon) \\ &\ge \mathbb{P}(X_n > \delta, |Y_{n+1} - 1| > \epsilon) \\ &= \mathbb{P}(X_n > \delta)\mathbb{P}(|Y_{n+1} - 1| > \epsilon). \end{aligned}$$

Here, $\mathbb{P}(|X_{n+1} - X_n| > \delta\epsilon)$ converges to 0 since X_n converges a.s. Since $\mathbb{P}(|Y_{n+1} - 1| > \epsilon) > 0$ for small ϵ, this implies that $\mathbb{P}(X_n > \delta) \to 0$. That is, $X_n \xrightarrow{p} 0$. Therefore $X = 0$ a.s.

(ii) By applying Jensen's inequality to $\log(Y \vee \delta)$, we have $\mathbb{E}\log(Y \vee \delta) \le \log \mathbb{E}[Y \vee \delta]$. Letting $\delta \to 0$, we have $\mathbb{E}\log Y \le \log \mathbb{E}Y = 0$. By the strong law of large numbers, it follows that

$$\frac{1}{n}\log X_n = \frac{1}{n}\sum_{m \le n} \log Y_m \to \mathbb{E}\log Y \le 0.$$

(Note that this holds even when $\mathbb{E}\log Y = -\infty$.)

Exercise 4.2.7. Confer Proposition 3.1 in Chapter 5 of Complex Analysis by Stein and Shakarchi.

Exercise 4.2.8. Observe that, from the given inequality,

$$\mathbb{E}\left[\frac{X_{n+1}}{\prod_{m=1}^{n}(1 + Y_m)} \middle| \mathcal{F}_n \right] \le \frac{X_n}{\prod_{m=1}^{n-1}(1 + Y_m)}.$$

Thus the random variable

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

which is also adapted and L^1, is a nonnegative supermartingale, and thus converges a.s. Since $\prod_{m=1}^{n-1}(1 + Y_m)$ converges a.s. by Exercise 4.2.7, it follows that X_n converges a.s.

Exercise 4.2.9. (i) On $\{N > n\}$,

$$\begin{aligned}\mathbb{E}[Y_{n+1}\mathbf{1}_{N>n}|\mathcal{F}_n] &= \mathbb{E}[X^1_{n+1}\mathbf{1}_{N>n+1} + X^2_{n+1}\mathbf{1}_{n<N\leq n+1}|\mathcal{F}_n] \\ &\leq \mathbb{E}[X^1_{n+1}\mathbf{1}_{N>n+1}|\mathcal{F}_n] + \mathbb{E}[X^1_{n+1}\mathbf{1}_{n<N\leq n+1}|\mathcal{F}_n] \\ &= \mathbb{E}[X^1_{n+1}\mathbf{1}_{N>n}|\mathcal{F}_n] \\ &= \mathbb{E}[X^1_{n+1}|\mathcal{F}_n]\mathbf{1}_{N>n} \\ &= X^1_n\mathbf{1}_{N>n}.\end{aligned}$$

Meanwhile, on $\{N \leq n\}$,

$$\begin{aligned}\mathbb{E}[Y_{n+1}\mathbf{1}_{N\leq n}|\mathcal{F}_n] &= \mathbb{E}[X^2_{n+1}\mathbf{1}_{N\leq n}|\mathcal{F}_n] \\ &= \mathbb{E}[X^2_{n+1}|\mathcal{F}_n]\mathbf{1}_{N\leq n} \\ &\leq X^2_n\mathbf{1}_{N\leq n}.\end{aligned}$$

Thus $\mathbb{E}[Y_{n+1}|\mathcal{F}_n] \leq X^1_n\mathbf{1}_{N>n} + X^2_n\mathbf{1}_{N\leq n} = Y_n$, and Y_n is a supermartingale.

(ii) Use the similar argument as the above.

Exercise 4.2.10. content

4.3. **Examples.**

Exercise 4.3.1. Let $X_0 = 0$. Define X_{n+1} as follows: If $X_n = 0$, then $X_{n+1} = 1$ or -1 with probability $1/2$ each. If $X_n \neq 0$, then $X_{n+1} = 0$ or n^2X_n with probability $1 - 1/n^2$ and $1/n^2$ respectively. Note that $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$. i.e. X_n is indeed a martingale. Note also that, since $\sum_n 1/n^2 < \infty$, by Borel-Cantelli, $\mathbb{P}(X_{n+1} = n^2X_n \text{ i.o.}|X_n \neq 0) = 0$. That is, given $X_n \neq 0$, $X_{n+1} = 0$ except for finitely many times. Thus $\mathbb{P}(X_n = a \text{ i.o.}) = 1$ for $a = -1, 0, 1$.

Exercise 4.3.2.

Exercise 4.3.3. Observe that from the given inequality,

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

Thus the random variable

$$Z_n = X_n - \sum_{k=1}^{n-1} Y_k,$$

which is also adapted and L^1, is a supermartingale. Let $N_M = \inf\{\sum_{k=1}^n Y_k > M\}$. Then $Z_{N_M\wedge n}$ is also a supermartingale. Since $Z_{N_M\wedge n} + M = X_{N_M\wedge n} + M - \sum_{k=1}^{N_M\wedge n} Y_k \geq 0$, it is a nonnegative supermartingale, and thus converges a.s. to an L^1 limit. That is, Z_n converges to an L^1 limit on the event $\{N_M = \infty\}$. Since $\sum Y_k < \infty$ a.s., if we let $M \to \infty$, $\{N_M = \infty\} \uparrow \Omega$ up to a measure zero set. Therefore, Z_n converges a.s. to an L^1 limit Z_∞ and thus $X_n = Z_n + \sum_{k=1}^{n-1} Y_k \overset{\text{a.s.}}{\to} Z_\infty + \sum Y_k < \infty$.

Exercise 4.3.4. Let $(A_m)_m$ be a sequence of independent events with $\mathbb{P}(A_m) = p_m$. By the Borel-Cantelli lemma, we know that $\sum p_m = \infty$ if and only if $\mathbb{P}(A_m \text{ i.o.}) = 1$. Thus we show that $\mathbb{P}(A_m \text{ i.o.}) = 1$ if and only if $\prod(1 - p_m) = 0$. Suppose first that $\mathbb{P}(A_m \text{ i.o.}) = 1$. Then

$$\begin{aligned} 0 &= \mathbb{P}(A_m^c \text{ eventually}) \\ &= \lim_{n\to\infty} \mathbb{P}(\bigcap_{m\geq n} A_m^c) \\ &= \lim_{n\to\infty} \prod_{m\geq n} (1 - p_m). \end{aligned}$$

Note that

$$\prod_{m=1}^{\infty} (1 - p_m) \leq \prod_{m\geq n} (1 - p_m) \to 0$$

and thus $\prod(1 - p_m) = 0$. Now for the other direction, suppose that $\prod(1 - p_m) = 0$. Since $p_m < 1$, $\prod_{1\leq m\leq n}(1 - p_m) \neq 0$, and thus it follows that

$$0 = \prod_{m\geq n+1} (1 - p_m) = \mathbb{P}(\bigcap_{m\geq n+1} A_m^c)$$

for all n. Letting $n \to \infty$, we get $\mathbb{P}(A_m^c \text{ eventually}) = 0$ and thus $\mathbb{P}(A_m \text{ i.o.}) = 1$.

Exercise 4.3.5. Suppose first that $\mathbb{P}(A_n | \cap_{m=1}^{n-1} A_m^c) = 1$ for some n. This means that $A_n \supset \cap_{m=1}^{n-1} A_m^c$ up to a measure zero set. That is, $\cap_{m=1}^{n} A_m^c = \cap_{m=1}^{n-1} A_m^c \setminus A_n$ is measure zero. Then we immediately have $\mathbb{P}(\cap_{m=1}^{\infty} A_m^c) = 0$. Thus we may assume that $\mathbb{P}(A_n | \cap_{m=1}^{n-1} A_m^c) < 1$ for all n. By Exercise 4.3.4, it follows that

$$\begin{aligned} 0 &= \prod_{n=2}^{\infty} \mathbb{P}(A_n^c | \cap_{m=1}^{n-1} A_m^c) \\ &= \prod_{n=2}^{\infty} \frac{\mathbb{P}(\cap_{m=1}^{n} A_m^c)}{\mathbb{P}(\cap_{m=1}^{n-1} A_m^c)} \\ &= \frac{\mathbb{P}(\cap_{m=1}^{\infty} A_m^c)}{\mathbb{P}(A_1^c)} \end{aligned}$$

and thus $\mathbb{P}(\cap_{m=1}^{\infty} A_m^c) = 0$. (Note that $\mathbb{P}(A_1^c) \neq 0$ from the above assumption.)

Exercise 4.3.6. (i) Note first that X_n is $\mathcal{F}_n$-measurable and in L^1. To see it is a martingale, we show that

$$\int_{I_{k,n}} X_{n+1}\, d\nu = \int_{I_{k,n}} X_n\, d\nu$$

for all k. If $\nu(I_{k,n}) = 0$, then $X_n \mathbf{1}_{I_{k,n}} = 0$ and $X_{n+1}\mathbf{1}_{I_{k,n}} = 0$, and thus the above holds. Now assume $\nu(I_{k,n}) \neq 0$ and suppose that $I_{k,n} = I_{p+1,n+1} \cup \cdots \cup I_{q,n+1} \cup I_{q+1,n+1} \cup \cdots \cup I_{r,n+1}$ where $\nu(I_{p+1,n+1}), \cdots, \nu(I_{q,n+1})$ are nonzero and $\nu(I_{q+1,n+1}) = \cdots = \nu(I_{r,n+1}) =$

0. Then we have

$$\begin{aligned}\int_{I_{k,n}} X_{n+1}\,d\nu &= \mu(I_{p+1,n+1}) + \cdots + \mu(I_{q,n+1}) \\ &= \mu(I_{p+1,n+1}) + \cdots + \mu(I_{q,n+1}) + \mu(I_{q+1,n+1}) + \cdots + \mu(I_{r,n+1}) \\ &= \mu(I_{k,n}) \\ &= \int_{I_{k,n}} X_n\,d\nu.\end{aligned}$$

Thus X_n is a martingale.

(ii) Again, note that X_n is $\mathcal{F}_n$-measurable and in L^1. Here, we show that

$$\int_{I_{k,n}} X_{n+1}\,d\nu \le \int_{I_{k,n}} X_n\,d\nu$$

for all k. If $\nu(I_{k,n}) = 0$, then the inequality trivially holds as the above. Assume $\nu(I_{k,n}) \neq 0$. With the same notation as the above, we have

$$\begin{aligned}\int_{I_{k,n}} X_{n+1}\,d\nu &= \mu(I_{p+1,n+1}) + \cdots + \mu(I_{q,n+1}) \\ &\le \mu(I_{p+1,n+1}) + \cdots + \mu(I_{q,n+1}) + \mu(I_{q+1,n+1}) + \cdots + \mu(I_{r,n+1}) \\ &= \mu(I_{k,n}) \\ &= \int_{I_{k,n}} X_n\,d\nu.\end{aligned}$$

Exercise 4.3.7. We first assume that μ and ν are finite measures. Let $\mathcal{F}_n = \sigma(A_m \colon 1 \le m \le n)$ and let μ_n and ν_n be restrictions of μ and ν to $\mathcal{F}_n$ respectively. Obviously $\mu_n \ll \nu_n$. Since $\mathcal{F}_n$ is finitely generated for each n, there exists a family of partitions $(I_{k,n})_{k,n}$ such that $\mathcal{F}_n = \sigma(I_{k,n} \colon 1 \le k \le K_n)$ and $\{I_{k,n+1} \colon 1 \le k \le K_{n+1}\}$ is a refinement of $\{I_{k,n} \colon 1 \le k \le K_n\}$. (Check this.) Define X_n as in Example 4.3.7. Then by definition, X_n satisfies

$$(*) \qquad \mu_n(A) = \int_A X_n\,d\nu_n$$

for $A \in \mathcal{F}_n$. (We don't define X_n as the Radon-Nikodym derivative because we are in the way of proving the Radon-Nikodym theorem.) Moreover, as we saw in Exercise 4.3.6, X_n is a martingale, and in particular it is a nonnegative supermartingale. By Theorem 4.2.12, $X_n \overset{\text{a.s.}}{\to} X$ where $X \in L^1(\nu)$. Now by the proof of Theorem 4.3.5,

$$\mu(A) = \int_A X\,d\nu + \mu(A \cap \{X = \infty\}).$$

(Even though X_n is not a Radon-Nikodym derivative, it plays the same role as in the proof of Theorem 4.3.5 because of $(*)$ above. Thus the proof works.) Since $X \in L^1(\nu)$, $\nu(\{X = \infty\}) = 0$ and by the absolute continuity $\mu(\{X = \infty\}) = 0$. Therefore,

$$\mu(A) = \int_A X\,d\nu.$$

Finally, note that X is measurable with respect to $\mathcal{F} = \sigma(A_n : n \in \mathbb{N})$. This proves the Radon-Nikodym theorem for finite measures.

Now, let μ and ν be σ-finite. Let $\Omega_k \uparrow \Omega$ be satisfying $\mu(\Omega_k) < \infty$ and $\nu(\Omega_k) < \infty$ for all k. (Check that we can always find such Ω_k's.) Define $\mu^{(k)}(\cdot) = \mu(\cdot \cap \Omega_k)$ and $\nu^{(k)}(\cdot) = \nu(\cdot \cap \Omega_k)$. Then $\mu^{(k)}$ and $\nu^{(k)}$ are finite measures. By the above, we have

$$\mu^{(k)}(A) = \int_A X^{(k)}\, d\nu^{(k)}$$

Where $X^{(k)}$ is $\mathcal{F}$-measurable and supported on Ω_k. Define X to be equal a.s. to X_k on Ω_k. Note that for $j < k$,

$$\mu^{(j)}(A) = \mu^{(k)}(A \cap \Omega_j) = \int_{A \cap \Omega_j} X^{(k)}\, d\nu^{(k)} = \int_A X^{(k)}\, d\nu^{(j)}.$$

Because of the uniqueness of (almost sure) limit, $X^{(k)} = X^{(j)}$ on Ω_j. Thus X is well-defined. Letting $k \to \infty$ and using the continuity of measures and the definition of Lebesgue integral,

$$\mu(A) = \lim_{k\to\infty} \mu^{(k)}(A) = \lim_{k\to\infty} \int_A X\, d\nu^{(k)} = \int_A X\, d\nu.$$

Finally it is obvious that X is $\mathcal{F}$-measurable.

Exercise 4.3.8. (i) Note that

$$\begin{aligned} 1 - \alpha_n = F_n(0) &= \int_{(-\infty,0]} q_n(x)\, dG_n(x) = q_n(0)(1-\beta_n), \\ 1 = F_n(1) &= \int_{(-\infty,1]} q_n(x)\, dG_n(x) = q_n(0)(1-\beta_n) + q_n(1)\beta_n. \end{aligned}$$

Thus $q_n(0) = (1-\alpha_n)/(1-\beta_n)$ and $q_n(1) = \alpha_n/\beta_n$. By Theorem 4.3.8, $\mu \ll \nu$ if and only if

$$\prod_{m=1}^{\infty} \int \sqrt{q_m}\, dG_m > 0.$$

Note that

$$\begin{aligned} \prod_{m=1}^{\infty} \int \sqrt{q_m}\, dG_m &= \prod_{m=1}^{\infty} \left(\sqrt{\frac{1-\alpha_m}{1-\beta_m}}(1-\beta_m) + \sqrt{\frac{\alpha_m}{\beta_m}}\beta_m \right) \\ &= \prod_{m=1}^{\infty} \left(\sqrt{(1-\alpha_m)(1-\beta_m)} + \sqrt{\alpha_m\beta_m} \right). \end{aligned}$$

Thus the condition becomes

$$\prod_{m=1}^{\infty} \left(\sqrt{(1-\alpha_m)(1-\beta_m)} + \sqrt{\alpha_m\beta_m} \right) > 0.$$

(ii)

Exercise 4.3.9. By the Borel-Cantelli lemma, $\sum \alpha_n < \infty$ implies that $\xi_n = 0$ eventually μ-a.s. and $\sum \beta_n = \infty$ implies that $\xi_n = 1$ infinitely often ν-a.s. Since $\{\xi_n = 0 \text{ eventually}\} \cap \{\xi_n = 1 \text{ i.o.}\} = \varnothing$, we have $\mu \perp \nu$.

Exercise 4.3.10. content

Exercise 4.3.11. content

Exercise 4.3.12. content

Exercise 4.3.13. content

4.4. **Doob's Inequality, Convergence in L^p, $p > 1$.**

Exercise 4.4.1. Since X is a submartingale, $\mathbb{E}[X_k \mathbf{1}_{N=j} | \mathcal{F}_j] = \mathbb{E}[X_k | \mathcal{F}_j] \mathbf{1}_{N=j} \geq X_j \mathbf{1}_{N=j}$. Now take the expectation of both sides.

Exercise 4.4.2. On $\{M = N\}$, the result is trivial. On $\{M < N\}$, let $K_n = \mathbf{1}_{M < n \leq N}$. Then K is predictable. Note that $(K \cdot X)_n = X_{N \wedge n} - X_{M \wedge n}$. Since $(K \cdot X)$ is a submartingale, from $\mathbb{E}(K \cdot X)_0 \leq \mathbb{E}(K \cdot X)_k$, we have $\mathbb{E}X_M \leq \mathbb{E}X_N$.
(Another proof) Let $Y_n = X_{N \wedge n}$. Then Y is a submartingale. By Theorem 4.4.1, $\mathbb{E}X_M = \mathbb{E}Y_M \leq \mathbb{E}Y_k = \mathbb{E}X_N$.

Exercise 4.4.3. We first observe that $A \in \mathcal{F}_N$. To see this, note that $\{N \leq n\} \cap A = \{N \leq n\} \cap \{M \leq n\} \cap A$, and that $\{M \leq n\} \cap A \in \mathcal{F}_n$ and $\{N \leq n\} \in \mathcal{F}_n$. Thus $\{N \leq n\} \cap A \in \mathcal{F}_n$, meaning that $A \in \mathcal{F}_N$. Now consider $\{L \leq n\}$. Observe that

$$\begin{aligned} \{L \leq n\} &= (\{L \leq n\} \cap A) \cup (\{L \leq n\} \cap A^c) \\ &= (\{M \leq n\} \cap A) \cup (\{N \leq n\} \cap A^c) \in \mathcal{F}_n \end{aligned}$$

since $A \in \mathcal{F}_M$ and $A \in \mathcal{F}_N$. Thus L is a stopping time.

Exercise 4.4.4. Let $A \in \mathcal{F}_M$ and L be the stopping time defined as above. Since $L \leq N$, by Exercise 4.4.2, $\mathbb{E}X_L \leq \mathbb{E}X_N$. Note that $\mathbb{E}X_L = \mathbb{E}X_L \mathbf{1}_A + \mathbb{E}X_L \mathbf{1}_{A^c} = \mathbb{E}X_M \mathbf{1}_A + \mathbb{E}X_N \mathbf{1}_{A^c}$. Plugging this into the inequality, we have $\mathbb{E}X_N \geq \mathbb{E}X_M \mathbf{1}_A + \mathbb{E}X_N \mathbf{1}_{A^c}$, which leads to $\mathbb{E}X_N \mathbf{1}_A \geq \mathbb{E}X_M \mathbf{1}_A$. Since $A \in \mathcal{F}_M$ was arbitrary, it follows that $\mathbb{E}[X_N | \mathcal{F}_M] \geq X_M$.

Exercise 4.4.5. We show that $\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]\mathbb{E}[Y|\mathcal{F}]] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}]^2]$. Since $\mathcal{F} \subset \mathcal{G}$,

$$\begin{aligned} \mathbb{E}[\mathbb{E}[Y|\mathcal{G}]\mathbb{E}[Y|\mathcal{F}]] &= \mathbb{E}[\mathbb{E}[Y\mathbb{E}[Y|\mathcal{F}]|\mathcal{G}]] \\ &= \mathbb{E}[Y\mathbb{E}[Y|\mathcal{F}]] \\ &= \mathbb{E}[\mathbb{E}[Y\mathbb{E}[Y|\mathcal{F}]|\mathcal{F}]] \\ &= \mathbb{E}[\mathbb{E}[Y|\mathcal{F}]^2]. \end{aligned}$$

This proves the result.

Exercise 4.4.6. Let $A = \{\max_{1\le m\le n} |S_m| > x\}$, and let $N = \inf\{m \ge 0\colon |S_m| > x\}$ and $T = N \wedge n$. Then N and T are stopping times. Since $S_n^2 - s_n^2$ is a martingale, by Exercise 4.4.1,

$$\begin{aligned}
0 = \mathbb{E}[S_0^2 - s_0^2] &= \mathbb{E}[S_T^2 - s_T^2] \\
&= \mathbb{E}[(S_N^2 - s_N^2)\mathbf{1}_A] + \mathbb{E}[(S_n^2 - s_n^2)\mathbf{1}_{A^c}] \\
&\le (x+K)^2\mathbb{P}(A) + (x^2 - s_n^2)\mathbb{P}(A^c).
\end{aligned}$$

Thus we have

$$0 \le (x+K)^2 - (x+K)^2\mathbb{P}(A^c) + (x^2 - s_n^2)\mathbb{P}(A^c)$$

and it follows that

$$\mathbb{P}(A^c) \le \frac{(x+K)^2}{s_n^2 + (x+K)^2 - x^2} \le \frac{(x+K)^2}{s_n^2}.$$

Exercise 4.4.7. Let $c \ge -\lambda$. Since $(X_n + c)^2$ is a nonnegative submartingale, by Doob's inequality,

$$\begin{aligned}
\mathbb{P}(\max_{1\le m\le n} X_m \ge \lambda) &= \mathbb{P}(\max_{1\le m\le n} (X_m + c)^2 \ge (\lambda + c)^2) \\
&\le \frac{\mathbb{E}(X_n + c)^2}{(\lambda + c)^2} \\
&= \frac{\mathbb{E}X_n^2 + c^2}{(\lambda + c)^2}
\end{aligned}$$

Optimizing over $c \ge -\lambda$, the right-hand-side is minimized when $c = \mathbb{E}X_n^2/\lambda$. Plugging this c into the inequality, the result follows.

Exercise 4.4.8. (i)

$$\begin{aligned}
\mathbb{E}(\bar{X}_n \wedge M) &= \int_0^\infty \mathbb{P}(\bar{X}_n \wedge M > x)\,dx \\
&\le \int_0^1 dx + \int_1^\infty \mathbb{P}(\bar{X}_n \wedge M > x)\,dx \\
&\le 1 + \int_1^\infty \frac{1}{x}\mathbb{E}(\bar{X}_n^+ \mathbf{1}_{\bar{X}_n \wedge M > x})\,dx \\
&= 1 + \int_1^\infty \int \frac{1}{x}\bar{X}_n^+ \mathbf{1}_{\bar{X}_n \wedge M > x}\,d\mathbb{P}\,dx \\
&= 1 + \int \int_1^{\bar{X}_n \wedge M} \frac{1}{x}\bar{X}_n^+\,dx\,d\mathbb{P} \\
&= 1 + \int \bar{X}_n^+ \log(\bar{X}_n \wedge M)\,d\mathbb{P}
\end{aligned}$$

(ii) Let $f(a,b) = (a\log b - a\log a)/b$. From $\nabla f = 0$ we get $\log b = \log a + 1$, or simply $b = ea$. Thus the maximum of f is

$$f(a, ea) = 1/e$$

which proves the result.

Exercise 4.4.9. From $\mathbb{E}X_n < \infty$ and $\mathbb{E}Y_n < \infty$, all the expectations are finite. Note that $\mathbb{E}X_{m-1}Y_m = \mathbb{E}[\mathbb{E}[X_{m-1}Y_m|\mathcal{F}_{m-1}]] = \mathbb{E}[X_{m-1}\mathbb{E}[Y_m|\mathcal{F}_{m-1}]] = \mathbb{E}X_{m-1}Y_{m-1}$. Thus $\sum_{m=1}^n \mathbb{E}(X_m - X_{m-1})(Y_m - Y_{m-1}) = \sum_{m=1}^n (\mathbb{E}X_mY_m - \mathbb{E}X_{m-1}Y_{m-1}) = \mathbb{E}X_nY_n - \mathbb{E}X_0Y_0$.

Exercise 4.4.10. Note that for all n,

$$\begin{aligned}
\mathbb{E}X_n^2 &= \mathbb{E}(X_0 + \xi_1 + \cdots + \xi_n)^2 \\
&= \mathbb{E}X_0^2 + \sum_{m=1}^n \mathbb{E}\xi_m^2 + 2\sum_{m=1}^n \mathbb{E}X_0\xi_m + 2\sum_{m,k=1}^n \mathbb{E}\xi_m\xi_k \\
&= \mathbb{E}X_0^2 + \sum_{m=1}^n \mathbb{E}\xi_m^2 \\
&\le \mathbb{E}X_0^2 + \sum_{m=1}^\infty \mathbb{E}\xi_m^2 < \infty
\end{aligned}$$

where the third equality is due to Theorem 4.4.7. By Theorem 4.4.6, the result follows.

Exercise 4.4.11. For all n,

$$\mathbb{E}\frac{X_n^2}{b_n^2} = \frac{1}{b_n^2}\mathbb{E}X_0^2 + \frac{1}{b_n^2}\sum_{m=1}^n \mathbb{E}\xi_m^2.$$

By Kronecker's lemma (Theorem 2.5.9), $\sum_{m=1}^n \mathbb{E}\xi_m^2/b_m^2$ converges to 0. Also, since $\mathbb{E}X_0^2 < \infty$ and $b_n \uparrow \infty$, $\mathbb{E}X_0^2/b_n^2$ converges to 0. Therefore $\mathbb{E}X_n^2/b_n^2$ converges to 0, which implies that $\sup \mathbb{E}X_n^2/b_n^2 < \infty$. Since X_n/b_n is a submartingale, by Theorem 4.4.6 X_n/b_n a.s. and in L^2. As $X_n/b_n \to 0$ in L^2, $X_n/b_n \to 0$ a.s.

4.6. **Uniform Integrability, Convergence in L^1.**

Exercise 4.6.1. Let $\mathcal{F}_n = \sigma(Y_1, \cdots, Y_n)$ and define $X_n = \mathbb{E}[\theta|\mathcal{F}_n]$. Then X is a martingale with filtration $(\mathcal{F}_n)$. By Theorem 4.6.8, X_n converges a.s. to $\mathbb{E}[\theta|\mathcal{F}_\infty]$. Now it remains to show that $\theta = \mathbb{E}[\theta|\mathcal{F}_\infty]$. i.e. θ is $\mathcal{F}_\infty$-measurable. It follows by the strong law of large numbers that almost surely

$$\frac{Y_1 + \cdots + Y_n}{n} \to \theta$$

and thus θ is $\mathcal{F}_\infty$-measurable.

Exercise 4.6.2. Obviously X_n is $\mathcal{F}_n$-measurable and $X_n \in L^1$ for all n since $|X_n| \le K$ as f is Lipschitz. Note that

$$\mathbb{E}X_{n+1}\mathbf{1}_{I_{k,n}} = \mathbb{E}X_{n+1}\mathbf{1}_{I_{2k+1,n+1}} + \mathbb{E}X_{n+1}\mathbf{1}_{I_{2k,n+1}}$$

$$= \frac{f\left(\frac{2k+2}{2^{n+1}}\right) - f\left(\frac{2k+1}{2^{n+1}}\right)}{\frac{1}{2^{n+1}}} \frac{1}{2^{n+1}} + \frac{f\left(\frac{2k+1}{2^{n+1}}\right) - f\left(\frac{2k}{2^{n+1}}\right)}{\frac{1}{2^{n+1}}} \frac{1}{2^{n+1}}$$
$$= f\left(\frac{2k+2}{2^{n+1}}\right) - f\left(\frac{2k}{2^{n+1}}\right)$$

and that

$$\begin{aligned}\mathbb{E}X_n \mathbf{1}_{I_{k,n}} &= \frac{f\left(\frac{k+1}{2^n}\right) - f\left(\frac{k}{2^n}\right)}{\frac{1}{2^n}} \frac{1}{2^n} \\ &= f\left(\frac{2k+2}{2^{n+1}}\right) - f\left(\frac{2k}{2^{n+1}}\right).\end{aligned}$$

Thus $\mathbb{E}X_{n+1}\mathbf{1}_{I_{k,n}} = \mathbb{E}X_n\mathbf{1}_{I_{k,n}}$ for all k, and by the definition of conditional expectation we have $X_n = \mathbb{E}[X_{n+1}|\mathcal{F}_n]$. That is, (X_n) is a martingale. Since $|X_n| \le K$ for all n, the family (X_n) is uniformly integrable. Therefore, by Theorem 4.6.7, X_n converges a.s. and in L^1 to an L^1 random variable X_∞. Let a_k and b_k be endpoints of some $I_{k,n}$. Observe that

$$f(b_k) - f(a_k) = \int_{a_k}^{b_k} X_n(\omega)\, d\omega.$$

By the dominated convergence, it follows that

$$f(b_k) - f(a_k) = \int_{a_k}^{b_k} X_\infty(\omega)\, d\omega$$

for all such points a_k and b_k. Since f is continuous and $|X_\infty| \le K$, letting $a_k \to a$ and $b_k \to b$ we have the last assertion.

Exercise 4.6.3. By Theorem 4.6.8, $\mathbb{E}[f|\mathcal{F}_n] \to \mathbb{E}[f|\mathcal{F}_\infty]$ a.s. and in L^1. Now it remains to show that f is $\mathcal{F}_\infty$-measurable, which is obvious.

Exercise 4.6.4. Let $\omega \in \{\limsup X_n < \infty\}$. We will show that $\omega \in D$. Let $M(\omega) = \limsup X_n(\omega) < \infty$. Then there exists $N(\omega)$ such that $X_n(\omega) < 2M(\omega)$ for all $n \ge N(\omega)$. From the given assumption, it follows that

$$\mathbb{P}(D|X_1, \cdots, X_n)(\omega) \ge \delta(x)$$

for all $n \ge N(\omega)$. Since $D \in \sigma(X_1, X_2, \cdots)$, by Theorem 4.6.9, the left-hand-side of the above display converges to $\mathbf{1}_D(\omega)$ for all $\omega \in \{\limsup X_n < \infty\}$ except on a measure zero set. That is, we have $\mathbf{1}_D \ge \delta(x) > 0$ for all $\omega \in \{\limsup X_n < \infty\} \setminus E$ where $\mathbb{P}(E) = 0$. It follows that $\mathbf{1}_D(\omega) = 1$ and thus $\omega \in D$. i.e. $\{\limsup X_n < \infty\} \setminus E \subset D$. The result now immediately follows.

Exercise 4.6.5. content

Exercise 4.6.6. Note that X_n is a martingale as $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n(\alpha+\beta X_n)+(1-X_n)\beta X_n = X_n$. Since $|X_n| \le 1$ for all n, it is a uniformly integrable martingale. Thus X_n converges to an L^1 random variable X almost surely (and in L^1).

To see $X \in \{0,1\}$, we first let $B_n = \{X_{n+1} = \alpha + \beta X_n\}$ and $B = \limsup B_n$. For $\omega \in B$, infinitely often,

$$\begin{aligned} X_{n+1}(\omega) - X_n(\omega) &= \alpha + \beta X_n(\omega) - X_n(\omega) \\ &= \alpha(1 - X_n(\omega)). \end{aligned}$$

Since X_n converges almost surely, $|X_{n+1}(\omega) - X_n(\omega)| < \epsilon$ for large enough n. That is, infinitely often we have

$$\alpha|1 - X_n(\omega)| < \epsilon.$$

Thus $X(\omega) = 1$ for almost every $\omega \in B$. Now consider $C = B^c = \liminf B_n^c$. For $\omega \in C$, eventually,

$$X_{n+1} = \beta X_n$$

which then immediately implies that $X(\omega) = 0$. Thus $X(\omega) = 0$ for almost every $\omega \in C$. This proves that $X \in \{0,1\}$

Finally, since X_n is a martingale, we have $\mathbb{E}X_0 = \mathbb{E}X_n$ and by the dominated convergence $\mathbb{E}X_n \to \mathbb{E}X$. Thus

$$\theta = \mathbb{E}X_0 = \mathbb{E}X = 1 \cdot \mathbb{P}(X = 1) + 0 \cdot \mathbb{P}(X = 0) = \mathbb{P}(X = 1).$$

Exercise 4.6.7. Note that

$$\begin{aligned} \mathbb{E}|\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_\infty]| &\le \mathbb{E}|\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_n]| + \mathbb{E}|\mathbb{E}[Y|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_\infty]| \\ &\le \mathbb{E}|Y_n - Y| + \mathbb{E}|\mathbb{E}[Y|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_\infty]| \end{aligned}$$

By Theorem 4.6.8, $\mathbb{E}[Y|\mathcal{F}_n] \to \mathbb{E}[Y|\mathcal{F}_\infty]$ in L^1 and $Y_n \to Y$ in L^1 by the assumption.

4.7. **Backwards Martingales.**

Exercise 4.7.1. By the L^p maximum inequality, for all $n \le 0$,

$$\mathbb{E}\left[\sup_{n \le m \le 0} |X_m|^p\right] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}|X_0|^p.$$

Letting $n \to -\infty$, it follows that $\sup_{m \le 0} |X_m| \in L^p$. Let $X_{-\infty}$ be the limit of the backward martingale X_n (which exists by Theorem 4.7.1). Noting that $|X_n - X_{-\infty}|^p \le (2\sup_{m \le 0} |X_m|)^p$, by the dominated convergence we have

$$\mathbb{E}|X_n - X_{-\infty}|^p \to 0.$$

Exercise 4.7.2. Note that

$$\begin{aligned} |\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y_{-\infty}|\mathcal{F}_{-\infty}]| &\le |\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y_{-\infty}|\mathcal{F}_n]| + |\mathbb{E}[Y_{-\infty}|\mathcal{F}_n] - \mathbb{E}[Y_{-\infty}|\mathcal{F}_{-\infty}]| \\ &\le \mathbb{E}[|Y_n - Y_{-\infty}||\mathcal{F}_n] + |\mathbb{E}[Y_{-\infty}|\mathcal{F}_n] - \mathbb{E}[Y_{-\infty}|\mathcal{F}_{-\infty}]|. \end{aligned}$$

By Theorem 4.7.3, $|\mathbb{E}[Y_{-\infty}|\mathcal{F}_n]-\mathbb{E}[Y_{-\infty}|\mathcal{F}_{-\infty}]| \to 0$. To see the convergence of the first term, let $W_N = \sup_{m,n\le N} |Y_m - Y_n|$. Then $W_N \le 2Z \in L^1$ and $W_N \to 0$. It follows that

$$\begin{aligned}\limsup_{n\to-\infty} \mathbb{E}[|Y_n - Y_{-\infty}||\mathcal{F}_n] &\le \limsup_{n\to-\infty} \mathbb{E}[W_N|\mathcal{F}_n] \\ &= \mathbb{E}[W_N|\mathcal{F}_{-\infty}] \\ &\to 0\end{aligned}$$

where the last convergence is due to the dominated convergence theorem.

Exercise 4.7.3. Let $p = \mathbb{P}(X_1 = 1, \cdots, X_m = 1)$. Because (X_n) is exchangeable, $\mathbb{P}(X_{n_1} = 1, \cdots, X_{n_m} = 1) = p$ for any subsequence (n_k). Thus $\mathbb{P}(S_n = m) = \binom{n}{m}p$. Similarly, $\mathbb{P}(X_1 = 1, \cdots, X_k = 1, X_{n_1} = 1, \cdots, X_{n_{m-k}} = 1) = p$ for any (n_k), and thus $\mathbb{P}(X_1 = 1, \cdots, X_k = 1, S_n = m) = \binom{n-k}{m-k}p$. Therefore,

$$\mathbb{P}(X_1 = 1, \cdots, X_k = 1|S_n = m) = \binom{n-k}{m-k} \Big/ \binom{n}{m} = \binom{n-k}{n-m} \Big/ \binom{n}{m}$$

Exercise 4.7.4. Note that from the exchangeability,

$$0 \le \mathbb{E}(X_1 + \cdots + X_n)^2 = n\mathbb{E}X_1^2 + n(n-1)\mathbb{E}X_1X_2$$

and thus

$$\mathbb{E}X_1X_2 \ge -\frac{\mathbb{E}X_1^2}{(n-1)} \to 0.$$

Exercise 4.7.5. Let $\varphi(x, y) = (x - y)^2$ and let

$$A_n(\varphi) = \binom{n}{2}^{-1} \sum_{1\le i<j\le n} \varphi(X_i, X_j).$$

Since $A_n(\varphi) \in \mathcal{E}_n$, we have

$$\begin{aligned}A_n(\varphi) &= \mathbb{E}[A_n(\varphi)|\mathcal{E}_n] \\ &= \binom{n}{2}^{-1} \sum_{1\le i<j\le n} \mathbb{E}[\varphi(X_i, X_j)|\mathcal{E}_n] \\ &= \mathbb{E}[\varphi(X_1, X_2)|\mathcal{E}_n].\end{aligned}$$

By Theorem 4.7.3, $\mathbb{E}[\varphi(X_1, X_2)|\mathcal{E}_n] \to \mathbb{E}[\varphi(X_1, X_2)|\mathcal{E}]$. (Note that this is possible because $\varphi(X_1, X_2)$ is integrable.) Now since $\mathcal{E}$ is trivial due to Example 4.7.6, it follows that

$$\begin{aligned}\mathbb{E}[\varphi(X_1, X_2)|\mathcal{E}] &= \mathbb{E}\varphi(X_1, X_2) \\ &= \mathbb{E}X_1^2 + \mathbb{E}X_2^2 - 2\mathbb{E}X_1\mathbb{E}X_2 \\ &= 2\sigma^2.\end{aligned}$$

4.8. **Optional Stopping Theorems.**

Exercise 4.8.1. Let $A \in \mathcal{F}_L$ and let $N = L\mathbf{1}_A + M\mathbf{1}_{A^c}$. Note that N is a stopping time according to Exercise 4.4.3 and that $N \le M$. Since $Y_{M\wedge n}$ is a uniformly integrable submartingale, applying Theorem 4.8.3 to this leads to

$$\mathbb{E}Y_0 \le \mathbb{E}Y_N \le \mathbb{E}Y_M.$$

Observe that $\mathbb{E}Y_N = \mathbb{E}Y_L\mathbf{1}_A + \mathbb{E}Y_M\mathbf{1}_{A^c}$. It follows that

$$\mathbb{E}Y_L\mathbf{1}_A \le \mathbb{E}Y_M\mathbf{1}_A.$$

Since $A \in \mathcal{F}_L$ was arbitrary, by the definition of conditional expectation,

$$Y_L \le \mathbb{E}[Y_M|\mathcal{F}_L]$$

and also $\mathbb{E}Y_L \le \mathbb{E}Y_M$.

Exercise 4.8.2. Let $N = \inf\{m\colon X_m > \lambda\}$. Since $X_N\mathbf{1}_{N<\infty} \ge \lambda$,

$$\lambda\mathbb{P}(N < \infty) \le \mathbb{E}X_N\mathbf{1}_{N<\infty} \le \mathbb{E}X_0$$

by Theorem 4.8.4. Observe that $\{N < \infty\} = \{\sup_m X_m > \lambda\}$.

Exercise 4.8.3. Let $X_n = S_n^2 - n\sigma^2$. Note that

$$\begin{aligned}\mathbb{E}|X_T| \le \mathbb{E}S_T^2 &= \mathbb{E}S_T^2\mathbf{1}_{T<\infty} + \mathbb{E}S_T^2\mathbf{1}_{T=\infty}\\ &\le \mathbb{E}(a+\xi)^2\mathbf{1}_{T<\infty} + a^2\\ &\le 2a^2 + \sigma^2 < \infty.\end{aligned}$$

To see

$$\mathbb{E}[|X_n\mathbf{1}_{T>n}|; |X_n\mathbf{1}_{T>n}| > a^2 + M] = 0$$

for all n, we observe that $|X_n\mathbf{1}_{T>n}| \le a^2\mathbf{1}_{T>n} \le a^2$ for all n and thus $\{|X_n\mathbf{1}_{T>n}| > a^2+M\} = \varnothing$. By Theorem 4.8.2, $X_{T\wedge n}$ is a uniformly integrable martingale. It follows that

$$0 = \mathbb{E}X_{T\wedge n} \to \mathbb{E}X_T = \mathbb{E}[S_T^2 - T\sigma^2].$$

If $\mathbb{E}T = \infty$, then the result is trivial. For $\mathbb{E}T < \infty$, we have $T < \infty$ a.s. and thus

$$\sigma^2\mathbb{E}T = \mathbb{E}S_T^2 \ge a^2.$$

Exercise 4.8.4. Let $X_n = S_n^2 - n\sigma^2$. Then X_n is a martingale. Since T is a stopping time with $\mathbb{E}T < \infty$, by the monotone convergence,

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

In particular, $\mathbb{E}S_{T\wedge n}^2 \le \sigma^2\mathbb{E}T$ for all n. i.e. $\sup_n \mathbb{E}S_{T\wedge n}^2 < \infty$. Since $S_{T\wedge n}$ is a martingale, this implies that $S_{T\wedge n}$ converges a.s. and in L^2 to S_T. Therefore, $\mathbb{E}S_{T\wedge n}^2 \to \mathbb{E}S_T^2$ and the result follows.

Exercise 4.8.5. (a) Let $A = p - q$ and $B = 1 - (p - q)^2$. Using the same argument as in Exercise 4.8.3, it follows that $(S_{V_0 \wedge n} - (V_0 \wedge n)A)^2 - (V_0 \wedge n)B$ is a uniformly integrable martingale. Thus we have

$$\begin{aligned} B\mathbb{E}V_0 &= \mathbb{E}(S_{V_0} - V_0 A)^2 \\ &= \mathbb{E}S_{V_0}^2 - 2A\mathbb{E}S_{V_0}V_0 + A^2\mathbb{E}V_0^2. \end{aligned}$$

Now since $p < 1/2$, $V_0 < \infty$ a.s. and $S_{V_0} = 0$. Thus the above display becomes

$$B\mathbb{E}V_0 = A^2\mathbb{E}V_0^2.$$

Therefore, the second moment of V_0 is (not the variance; probably a typo)

$$\mathbb{E}V_0^2 = \frac{B}{A^2}\mathbb{E}V_0 = \frac{1 - (p - q)^2}{(q - p)^3}x.$$

Exercise 4.8.6. (a) Let $\phi(\theta) = \mathbb{E}\exp(\theta\xi_i) = pe^{\theta} + qe^{-\theta}$ and $X_n = \exp(\theta S_n)/\phi(\theta)^n$. Then X_n is a martingale and thus

$$\begin{aligned} e^{\theta x} = \mathbb{E}X_0 &= \mathbb{E}X_{V_0 \wedge n} \\ &= \mathbb{E}\frac{\exp(\theta S_{V_0 \wedge n})}{\phi(\theta)^{V_0 \wedge n}}. \end{aligned}$$

Now note that, since $\theta \le 0$ and $S_{V_0 \wedge n} \ge 0$, $\exp(\theta S_{V_0 \wedge n}) \le 1$ and $\phi(\theta) \ge 1$. By the bounded convergence theorem, it follows that

$$\begin{aligned} e^{\theta x} &= \mathbb{E}\frac{\exp(\theta S_{V_0 \wedge n})}{\phi(\theta)^{V_0 \wedge n}} \\ &\to \mathbb{E}\frac{\exp(\theta S_{V_0})}{\phi(\theta)^{V_0}} \\ &= \mathbb{E}[\phi(\theta)^{-V_0}]. \end{aligned}$$

(b) From $\phi(\theta) = 1/s$, we get $pe^{2\theta} - e^{\theta}/s + q = 0$. Solving this with respect to e^{θ} and noting that $e^{\theta} \le 1$, we have

$$e^{\theta} = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps}.$$

Thus

$$\mathbb{E}[s^{V_0}] = \left(\frac{1 - \sqrt{1 - 4pqs^2}}{2ps}\right)^x.$$

Exercise 4.8.7. Note that $\mathbb{E}[S_{n+1}^4|\mathcal{F}_n] = S_n^4 + 6S_n^2 + 1$ and $\mathbb{E}[S_{n+1}^2|\mathcal{F}_n] = S_n^2 + 1$. From $\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = Y_n$, we get $(2b - 6)n + (b + c - 5) = 0$ and thus $b = 3$ and $c = 2$. Now from $\mathbb{E}Y_0 = \mathbb{E}Y_{T \wedge n}$,

$$0 = \mathbb{E}S_{T \wedge n}^4 - 6\mathbb{E}(T \wedge n)S_{T \wedge n}^2 + 3\mathbb{E}(T \wedge n)^2 + 2\mathbb{E}(T \wedge n).$$

Noting that $|S_{T\wedge n}| \le a$ and that $\mathbb{E}T < \infty$ (from the proof of Theorem 4.8.7), apply the DCT and MCT to get

$$0 = a^4 - 6a^2 \mathbb{E}T + 3\mathbb{E}T^2 + 2\mathbb{E}T.$$

Now $\mathbb{E}T = a^2$ leads to

$$\mathbb{E}T^2 = \frac{5a^4 - 2a^2}{3}.$$

Exercise 4.8.8. Note that $S_{\tau\wedge n} \ge a$, and hence $X_n \le e^{a\theta_0}$ for all n. Thus $X_{\tau\wedge n}$ is a uniformly integrable martingale, and by Theorem 4.8.2, $1 = \mathbb{E}X_0 = \mathbb{E}X_\tau$. Now note that

$$\begin{aligned} 1 = \mathbb{E}X_\tau &\ge \mathbb{E}[X_\tau; S_\tau \le a] \\ &\ge e^{a\theta_0}\mathbb{P}(S_\tau \le a) \end{aligned}$$

from which the result follows.

Exercise 4.8.9. By the same argument as in Exercise 4.8.8, $X_{T\wedge n}$ is a uniformly integrable martingale and we have $1 = \mathbb{E}X_0 = \mathbb{E}X_T$. Now observe that

$$\begin{aligned} 1 = \mathbb{E}X_T &= \mathbb{E}[X_T; T < \infty] + \mathbb{E}[X_T; T = \infty] \\ &= e^{a\theta_0}\mathbb{P}(T < \infty) + \mathbb{E}[X_\infty; T = \infty]. \end{aligned}$$

Here, $X_\infty = \exp(\theta_0 S_\infty) = 0$ since $S_\infty = \lim S_n = \infty$ a.s. by the strong law of large numbers. Therefore,

$$\mathbb{P}(T < \infty) = e^{-a\theta_0}.$$

Exercise 4.8.10. Note that $\phi(\theta) = \mathbb{E}\exp(\theta\xi) = (e^{-\theta} + e^{\theta} + e^{2\theta})/3$. Solving $\phi(\theta) = 1$ with $\theta \le 0$, we get $e^\theta = \sqrt{2} - 1$. By Exercise 4.8.9, $\mathbb{P}(T < \infty) = e^{\theta i} = (\sqrt{2} - 1)^i$.

Exercise 4.8.11. Let $\phi(\theta) = \mathbb{E}\exp(\theta\xi) = \exp((c-\mu)\theta + \sigma^2\theta^2/2)$. Note that $\theta_0 = -2(c-\mu)/\sigma^2 \le 0$ satisfies $\phi(\theta_0) = 1$. For this θ_0, $X_n = \exp(\theta_0 S_n)$ is a martingale. Now let $T = \inf\{S_n \le 0\}$. By the same argument as in Exercise 4.8.8, $X_{T\wedge n}$ is a uniformly integrable martingale. Thus

$$\begin{aligned} \mathbb{E}\exp(\theta_0 S_0) &= \mathbb{E}\exp(\theta_0 S_T) \\ &\ge \mathbb{E}[\exp(\theta_0 S_T); T < \infty] \\ &\ge \mathbb{P}(T < \infty). \end{aligned}$$

Now plugging $\theta_0 = -2(c-\mu)/\sigma^2$ into the last display, the result follows.

4.9. **Combinatorics of Simple Random Walk.**

Exercise 4.9.1. content

5. Markov Chains

5.1. **Examples.**

Exercise 5.1.1.

$$\mathbb{P}(X_{n+1}=j|X_n=i,X_{n-1}=i_{n-1},\cdots,X_0=i_0)=\begin{cases}1-\dfrac{i}{N} & \text{for } j=i+1,\\ \dfrac{i}{N} & \text{for } j=i.\end{cases}$$
$$=: p(i,j)$$

Exercise 5.1.2. Note that

$$\begin{aligned}\mathbb{P}(X_4=2|X_3=1,X_2=0,X_1=0)&=\frac{\mathbb{P}(X_4=2,X_3=1,X_2=0,X_1=0)}{\mathbb{P}(X_3=1,X_2=0,X_1=0)}\\&=\frac{\mathbb{P}(S_4=2,S_3=1,S_2=0,S_1=-1)}{\mathbb{P}(S_3=1,S_2=0,S_1=-1)}=\frac{1}{2},\end{aligned}$$

$$\begin{aligned}\mathbb{P}(X_4=2|X_3=1,X_2=1,X_1=1)&=\frac{\mathbb{P}(X_4=2,X_3=1,X_2=1,X_1=1)}{\mathbb{P}(X_3=1,X_2=1,X_1=1)}\\&=\frac{\mathbb{P}(S_4=2,S_3=1,S_2=0,S_1=1)}{\mathbb{P}(S_2=0,S_1=1)}=\frac{1}{4}.\end{aligned}$$

Thus X_n is not a Markov chain.

Exercise 5.1.3. Note that $X_i \in \sigma(\xi_n,\cdots,\xi_0)$ for $i \le n-1$. Since $X_{n+1}=(\xi_{n+1},\xi_{n+2})$, X_{n+1} is independent of $\{X_{n-1},\cdots,X_0\}$. Thus,

$$\mathbb{P}(X_{n+1}=j|X_n=i,X_{n-1}=i_{n-1},\cdots,X_0=i_0)=\mathbb{P}(X_{n+1}=j|X_n=i).$$

That is, X_n is a Markov chain. The transition probabilities are

p	HH	HT	TH	TT
HH	1/2	1/2	0	0
HT	0	0	1/2	1/2
TH	1/2	1/2	0	0
TT	0	0	1/2	1/2

and

p^2	HH	HT	TH	TT
HH	1/4	1/4	1/4	1/4
HT	1/4	1/4	1/4	1/4
TH	1/4	1/4	1/4	1/4
TT	1/4	1/4	1/4	1/4

Indeed, the distribution of $X_{n+2}=(\xi_{n+2},\xi_{n+3})$ is independent of $X_n=(\xi_n,\xi_{n+1})$.

Exercise 5.1.4. The transition probabilities are

	AA,AA	AA,Aa	AA,aa	Aa,Aa	Aa,aa	aa,aa
AA,AA	1	0	0	0	0	0
AA,Aa	1/4	1/2	0	1/4	0	0
AA,aa	0	0	0	1	0	0
Aa,Aa	1/16	1/4	1/8	1/4	1/4	1/16
Aa,aa	0	0	0	1/4	1/2	1/4
aa,aa	0	0	0	0	0	1

Exercise 5.1.5.

$$\mathbb{P}(X_{n+1}=j|X_n=i)=\begin{cases}\frac{m-i}{m}\times\frac{b-i}{m} & \text{for } j=i+1\\ \frac{m-i}{m}\times\frac{m-b+i}{m}+\frac{i}{m}\times\frac{b-i}{m} & \text{for } j=i\\ \frac{i}{m}\times\frac{m-b+i}{m} & \text{for } j=i-1\end{cases}$$

Exercise 5.1.6. Let $K_n=|\{1\le i\le n\colon X_i=1\}|$. Note that $S_n=K_n-(n-K_n)=2K_n-n$, or $K_n=(S_n+n)/2$. From the elementary definition of conditional probability,

$$\mathbb{P}(X_{n+1}=1|X_n,\cdots,X_1)=\frac{\int_0^1\theta^{K_n+1}(1-\theta)^{n-K_n}}{\int_0^1\theta^{K_n}(1-\theta)^{n-K_n}}.$$

Here, integrating by parts,

$$\int_0^1\theta^m(1-\theta)^k=(-1)^{2m}\int_0^1\frac{m!}{(k+1)\cdots(k+m)}(1-\theta)^{k+m}=\frac{m!k!}{(k+m+1)!}.$$

Thus,

$$\mathbb{P}(X_{n+1}=1|X_n,\cdots,X_1)=\frac{K_n+1}{n+2}=\frac{S_n+n+2}{2(n+2)}.$$

Now, note that

$$\begin{aligned}\mathbb{P}(S_{n+1}=s_{n+1}|S_n,\cdots,S_1)&=\begin{cases}\mathbb{P}(X_{n+1}=1|X_n,\cdots,X_1) & \text{for } s_{n+1}=S_n+1\\ 1-\mathbb{P}(X_{n+1}=1|X_n,\cdots,X_1) & \text{for } s_{n+1}=S_n-1\end{cases}\\ &=\begin{cases}\frac{S_n+n+2}{2(n+2)} & \text{for } s_{n+1}=S_n+1\\ 1-\frac{S_n+n+2}{2(n+2)} & \text{for } s_{n+1}=S_n-1.\end{cases}\end{aligned}$$

That is, the transition probability depends only on S_n. Therefore, S_n is a Markov chain.

5.2. **Construction, Markov Properties.**

Exercise 5.2.1.

$$\begin{aligned}\mathbb{P}_\mu(A\cap B|X_n)&=\mathbb{E}_\mu\left[\mathbb{E}_\mu[\mathbf{1}_A\mathbf{1}_B|\mathcal{F}_n]|X_n\right]\\ &=\mathbb{E}_\mu\left[\mathbf{1}_A\mathbb{E}_\mu[\mathbf{1}_B|\mathcal{F}_n]|X_n\right]\\ &\overset{(*)}{=}\mathbb{E}_\mu\left[\mathbf{1}_A\mathbb{E}_\mu[\mathbf{1}_B|X_n]|X_n\right]\\ &=\mathbb{E}_\mu\left[\mathbf{1}_A|X_n\right]\mathbb{E}_\mu[\mathbf{1}_B|X_n]\end{aligned}$$

Here, $(*)$ is from the Markov property.

Exercise 5.2.2. Let $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$, $\mathcal{F}_\infty = \sigma(X_1, X_2, \cdots)$. Note that $\bigcup_{m \ge n+1}\{X_m \in B_m\} \in \mathcal{F}_\infty$. By Lévy's 0-1 law, on $\{X_n \in A_n\}$,

$$\begin{aligned}\delta \le \mathbb{P}(\bigcup_{m\ge n+1}\{X_m \in B_m\}|X_n) &\overset{(*)}{=} \mathbb{P}(\bigcup_{m\ge n+1}\{X_m \in B_m\}|\mathcal{F}_n) \\ &\to \mathbf{1}_{\bigcup_{m\ge n+1}\{X_m \in B_m\}}\end{aligned}$$

where $(*)$ is from the Markov property. That is, for $\omega \in \{X_n \in A_n\}$, $\mathbf{1}_{\bigcup_{m\ge n+1}\{X_m\in B_m\}}(\omega) = 1$, or in other words,

$$\{X_n \in A_n\} \subset \bigcup_{m\ge n+1}\{X_m \in B_m\}.$$

Thus,

$$\bigcup_{n\ge k}\{X_n \in A_n\} \subset \bigcup_{m\ge k+1}\{X_m \in B_m\}$$

and letting $k \to \infty$, $\{X_n \in A_n \text{ i.o.}\} \subset \{X_n \in B_n \text{ i.o.}\}$.

Exercise 5.2.3. Let $A_\delta = \{x\colon \mathbb{P}_x(D) \ge \delta\}$. Then obviously $\mathbb{P}_{X_n}(D) \ge \delta$ on $\{X_n \in A_\delta\}$. Note that

$$\mathbb{P}_{X_n}(D) = \mathbb{P}\left(\bigcup_{m\ge n+1}\{X_m = a\}|X_n\right) \ge \delta$$

on $\{X_n \in A_\delta\}$. By the previous exercise,

$$\mathbb{P}(\{X_n \in A_\delta \text{ i.o.}\} - \{X_n = a \text{ i.o.}\}) = 0.$$

From $\{X_n = a \text{ i.o.}\} \subset D$, we have

$$\{X_n \in A_\delta \text{ i.o.}\} \subset \{X_n = a \text{ i.o.}\} \subset D.$$

This implies that $\mathbb{P}(X_n \in A_\delta^c \text{ eventually}) = 1$ on D^c. That is, on D^c, $\mathbb{P}_{X_n}(D) \le \delta$ eventually almost surely. Since δ is arbitrary, we have the desired result.

Exercise 5.2.4.

$$\begin{aligned}p^n(x,y) = \mathbb{P}_x(X_n = y) &= \sum_{m=1}^n \mathbb{P}_x(T_y = m)\mathbb{P}_x(X_n = y|T_y = m) \\ &= \sum_{m=1}^n \mathbb{P}_x(T_y = m)\mathbb{P}_x(X_n = y|X_1 \ne y, \cdots, X_{m-1} \ne y, X_m = y) \\ &\overset{(*)}{=} \sum_{m=1}^n \mathbb{P}_x(T_y = m)\mathbb{P}_x(X_n = y|X_m = y) \\ &= \sum_{m=1}^n \mathbb{P}_x(T_y = m)p^{n-m}(y,y)\end{aligned}$$

where $(*)$ is by the Markov property.

Exercise 5.2.5. Let $T^{(k)} = \inf\{i \geq k \colon X_i = x\} \geq k$. Then, similarly to the previous exercise,

$$\mathbb{P}_x(X_m = x) = \sum_{j=k}^{m} \mathbb{P}_x(T^{(k)} = j)\mathbb{P}_x(X_{m-j} = x).$$

Summing over m and applying Fubini,

$$\begin{aligned}
\sum_{m=k}^{n+k} \mathbb{P}_x(X_m = x) &= \sum_{m=k}^{n+k}\sum_{j=k}^{m} \mathbb{P}_x(T^{(k)} = j)\mathbb{P}_x(X_{m-j} = x) \\
&= \sum_{j=k}^{n+k}\sum_{m=j}^{n+k} \mathbb{P}_x(T^{(k)} = j)\mathbb{P}_x(X_{m-j} = x) \\
&= \sum_{j=k}^{n+k} \mathbb{P}_x(T^{(k)} = j) \sum_{m=0}^{n-(j-k)} \mathbb{P}_x(X_m = x) \\
&\leq \sum_{j=k}^{n+k} \mathbb{P}_x(T^{(k)} = j) \sum_{m=0}^{n} \mathbb{P}_x(X_m = x) \\
&= \mathbb{P}_x(T^{(k)} \leq n+k) \sum_{m=0}^{n} \mathbb{P}_x(X_m = x) \\
&\leq \sum_{m=0}^{n} \mathbb{P}_x(X_m = x).
\end{aligned}$$

Exercise 5.2.6. Since $\mathbb{P}_x(T_C < \infty) > 0$, there exist n_x and ϵ_x such that $\mathbb{P}_x(T_C \leq n_x) = \epsilon_x > 0$. Let $N = \max_{s \in S-C} n_x$ and $\epsilon = \min_{s\in S-C} \epsilon_x$. Then for all $y \in S - C$, $\mathbb{P}_y(T_C > N) \leq 1 - \epsilon$, and thus the desired result holds for $k = 1$. Now we proceed by induction. Note that

$$\begin{aligned}
\mathbb{P}_y(T_C > kN) &= \mathbb{P}_y(X_1 \notin C, \cdots, X_{kN} \notin C) \\
&= \mathbb{P}_y(X_{N+1} \notin C, \cdots, X_{kN} \notin C | X_1 \notin C, \cdots, X_N \notin C)\mathbb{P}_y(X_1 \notin C, \cdots, X_N \notin C) \\
&\leq (1-\epsilon)\mathbb{P}_y(X_{N+1} \notin C, \cdots, X_{kN} \notin C | X_N \notin C)
\end{aligned}$$

Here,

$$\begin{aligned}
\mathbb{P}_y(X_{N+1} \notin C, \cdots, X_{kN} \notin C | X_N \notin C) &= \frac{\mathbb{P}_y(X_N \notin C, \cdots, X_{kN} \notin C)}{\mathbb{P}_y(X_N \notin C)} \\
&= \frac{1}{\mathbb{P}_y(X_N \notin C)} \int_{\{X_N \notin C\}} \mathbf{1}_{\{X_{N+1}\notin C, \cdots, X_{kN} \notin C\}} \, d\mathbb{P}_y \\
&= \frac{1}{\mathbb{P}_y(X_N \notin C)} \int_{\{X_N \notin C\}} \mathbb{E}_y[\mathbf{1}_{\{X_{N+1}\notin C, \cdots, X_{kN} \notin C\}} | X_N] \, d\mathbb{P}_y \\
&= \frac{1}{\mathbb{P}_y(X_N \notin C)} \int_{\{X_N \notin C\}} \mathbb{E}_{X_N}[\mathbf{1}_{\{X_1\notin C, \cdots, X_{(k-1)N} \notin C\}}] \, d\mathbb{P}_y.
\end{aligned}$$

By the induction hypothesis, $\mathbb{E}_{X_N}[\mathbf{1}_{\{X_1\notin C, \cdots, X_{(k-1)N} \notin C\}}] \leq (1-\epsilon)^{k-1}$ on $\{X_N \notin C\}$ $\mathbb{P}_y$-almost surely. Thus the desired result follows for general k.

Exercise 5.2.7.

5.3. **Recurrence and Transience.**

5.4. **Recurrence of Random Walks.**

5.5. **Stationary Measures.**

5.6. **Asymptotic Behavior.**

5.7. **Periodicity, Tail σ-Field.**

5.8. **General State Space.**

6. Ergodic Theorems

6.1. **Definitions and Examples.**

Exercise 6.1.1. (i) Since $\varphi^{-1}\Omega \subset \Omega$ and $\varphi^{-1}\Omega$ has probability 1, $\varphi^{-1}\Omega = \Omega$ up to a null set. Thus $\Omega \in \mathcal{I}$.

(ii) Let $A \in \mathcal{I}$. Then up to a null set,

$$\varphi^{-1}(\Omega \setminus A) = \varphi^{-1}\Omega \setminus \varphi^{-1}A = \Omega \setminus A,$$

and thus $\Omega \setminus A \in \mathcal{I}$.

(iii) Let $A_j \in \Omega$ for $j \geq 1$. Then up to a null set,

$$\varphi^{-1}(\bigcup_j A_j) = \bigcup_j \varphi^{-1}A_j = \bigcup_j A_j,$$

and thus $\bigcup_j A_j \in \mathcal{I}$.

Now to see that $X \in \mathcal{I}$ if and only if $X \circ \varphi = X$ a.s., suppose first that $X \in \mathcal{I}$. Then up to a null set,

$$\{X \circ \varphi > a\} = \varphi^{-1}\{X > a\} = \{X > a\}$$

since $\{X > a\} \in \mathcal{I}$. Since this is true for arbitrary a, $X \circ \varphi = X$ a.s. (Confer the argument in Exercise 4.1.10.) For the other direction, suppose $X \circ \varphi = X$ a.s. Then up to a null set,

$$\{X > a\} = \{X \circ \varphi > a\} = \varphi^{-1}\{X > a\},$$

and thus $\{X > a\} \in \mathcal{I}$.

Exercise 6.1.2. (i)

$$\varphi^{-1}(B) = \varphi^{-1}(\bigcup_{n=0}^{\infty} \varphi^{-n}(A)) = \bigcup_{n=1}^{\infty} \varphi^{-n}(A) \subset B$$

(ii)

$$\varphi^{-1}(C) = \varphi^{-1}(\bigcap_{n=0}^{\infty} \varphi^{-n}(B)) = \bigcap_{n=1}^{\infty} \varphi^{-n}(B) = \bigcap_{n=1}^{\infty} (\varphi^{-n}(B) \cap B) = C$$

since $\varphi^{-n}(B) \cap B = \varphi^{-n}(B)$.

(iii) Suppose that A is almost invariant. Define $B = \bigcup_{n=0}^\infty \varphi^{-n}(A)$ and $C = \bigcap_{n=0}^\infty \varphi^{-n}(B)$. Then by the above, C is invariant in the strict sense. Note that

$$\begin{aligned} C &= \bigcap_{n=0}^\infty \varphi^{-n}(\bigcup_{k=0}^\infty \varphi^{-k}(A)) \\ &= \bigcap_{n=0}^\infty \bigcup_{k=0}^\infty \varphi^{-(n+k)}(A) \\ &= \bigcap_{n=0}^\infty \bigcup_{k\geq n} \varphi^{-k}(A) \\ &= \limsup_{n\to\infty} \varphi^{-n}(A). \end{aligned}$$

Thus both $C\backslash A = \limsup(\varphi^{-n}(A)\backslash A)$ and $A\backslash C = \liminf(A\backslash\varphi^{-n}(A))$ are measure-zero. Conversely, suppose that there exists C invariant in the strict sense with $\mathbb{P}(A\Delta C) = 0$. Then both $\varphi^{-1}(A \setminus C) = \varphi^{-1}(A) \setminus C$ and $\varphi^{-1}(C \setminus A) = C \setminus \varphi^{-1}(A)$ are measure-zero. It follows that both $\varphi^{-1}(A) \setminus A \subset (\varphi^{-1}(A) \setminus C) \cup (C \setminus A)$ and $A \setminus \varphi^{-1}(A) \subset (A \setminus C) \cup (C \setminus \varphi^{-1}(A))$ are measure-zero.

Exercise 6.1.3. (i) Suppose for the sake of contradiction that $x_n = x_m$ for some $n \neq m$. Then $(n-m)\theta = k \in \mathbb{N}$ and $\theta \in \mathbb{Q}$, which is a contradiction. Thus x_n are distinct in $[0,1)$, and hence for any N, $|x_m - x_n| < 1/N$ for some $m < n \leq N$. Now let $\epsilon > 0$ and $x \in [0,1)$. Fix $N > 1/\epsilon$. Let k be satisfying $x \in [k/N, (k+1)/N)$.

(ii)

(iii)

Exercise 6.1.4. Let $(Y_k)_{k\in\mathbb{Z}}$ be a sequence of random variables with joint distribution defined by

$$(Y_{-k}, \cdots, Y_k) \overset{d}{=} (X_0, \cdots, X_{2k}).$$

To see (Y_k) is well-defined, note that, by the stationarity of X,

$$(Y_{-k}, \cdots, Y_k) \overset{d}{=} (X_1, \cdots, X_{2k+1}) \overset{d}{=} (X_0, \cdots, X_{2k}).$$

Thus, distributions of $(Y_{-k}, \cdots, Y_k)_k$ are consistent, and by the Kolmogorov extension theorem, the sequence (Y_k) is well-defined. To check (Y_k) is stationary, observe that

$$(Y_0, \cdots, Y_k) \overset{d}{=} (X_k, \cdots, X_{2k}) \overset{d}{=} (X_{2n+k}, \cdots, X_{2n+2k}) \overset{d}{=} (Y_n, \cdots, Y_{n+k}).$$

Finally, note that $(X_0, \cdots, X_k) \overset{d}{=} (X_k, \cdots, X_{2k}) \overset{d}{=} (Y_0, \cdots, Y_k)$, and therefore (X_k) is indeed embedded in (Y_k).

Exercise 6.1.5. To see that Y_k is stationary, note that

$$\begin{aligned} (Y_0, \cdots, Y_k) &= (g((X_j)_{j\geq 0}), \cdots, g((X_{k+j})_{j\geq 0})) \\ &\overset{d}{=} (g((X_{n+j})_{j\geq 0}), \cdots, g((X_{n+k+j})_{j\geq 0})) \end{aligned}$$

$$= (Y_n, \cdots, Y_{n+k}).$$

To see that Y_k is ergodic, note that

$$Y_{k+1} = g(X_{k+1}, X_{k+2}, \cdots) = g(X_k \circ \varphi, X_{k+1} \circ \varphi, \cdots)) = Y_k \circ \varphi.$$

Thus, for $B \in \mathcal{B}(\mathbb{R}^{\{0,1,\cdots\}})$,

$$\begin{aligned} \{(Y_1, Y_2, \cdots) \in B\} &= \{(Y_0, Y_1, \cdots) \circ \varphi \in B\} \\ &= \varphi^{-1}\{(Y_0, Y_1, \cdots) \in B\} \end{aligned}$$

Thus, if we let $A = \{(Y_0, Y_1, \cdots) \in B\}$, then since X_n is ergodic (or in other words, since φ is ergodic), $A = \varphi^{-1}A$ implies that A is trivial. Therefore Y_n is also ergodic.

Exercise 6.1.6. Fix $N > n$. Let x_ν and y_ν be the smallest integer ≥ 0 that satisfy $\nu + x_\nu \equiv 0 \mod n$ and $\nu + N - y_\nu \equiv 0 \mod n$. Then x_ν and y_ν are uniformly distributed on $\{1, \cdots, n\}$. Note that

$$\begin{aligned} (Z_1, \cdots, Z_N) &= (Y_{\nu+1}, \cdots, Y_{\nu+N}) \\ &= (Y_{\nu+1}, \cdots, Y_{\nu+x_\nu}, \text{iid blocks of length } n, Y_{\nu+N-y_\nu+1}, \cdots, Y_{\nu+N}) \end{aligned}$$

and all the blocks that appear in the above (including the first and the last incomplete blocks) are independent. Now fix k and let $\tilde{x}_\nu$ and $\tilde{y}_\nu$ be the smallest integer ≥ 0 that satisfy $k + \nu + \tilde{x}_\nu \equiv 0 \mod n$ and $k + \nu + N - \tilde{y}_\nu \equiv 0 \mod n$. Then again $\tilde{x}_\nu$ and $\tilde{y}_\nu$ are uniformly distributed on $\{1, \cdots, n\}$. Note again that

$$\begin{aligned} (Z_{k+1}, \cdots, Z_{k+N}) &= (Y_{k+\nu+1}, \cdots, Y_{k+\nu+N}) \\ &= (Y_{k+\nu+1}, \cdots, Y_{k+\nu+\tilde{x}_\nu}, \text{iid blocks of length } n, Y_{k+\nu+N-\tilde{y}_\nu+1}, \cdots, Y_{k+\nu+N}). \end{aligned}$$

Since ν is independent of Y, so is $x_\nu, y_\nu, \tilde{x}_\nu$, and $\tilde{y}_\nu$. Thus, the distribution of $(Z_1, \cdots, Z_N)$ and $(Z_{k+1}, \cdots, Z_{k+N})$ are equal. To see that Z is ergodic, note that the tail σ-field $\mathcal{T}$ of Z is contained in the tail σ-field of blocks $(Y_{kn+1}, \cdots, Y_{kn+n})$. Since the blocks $(Y_{kn+1}, \cdots, Y_{kn+n})$ are iid, $\mathcal{T}$ is trivial. As the invariant σ-field is contained in $\mathcal{T}$, it is also trivial.

Exercise 6.1.7. Note first that

$$\mu([a, b]) = \frac{1}{\log 2} \int_a^b \frac{dx}{1+x} = \frac{1}{\log 2} \log \frac{1+b}{1+a}.$$

From the definition of φ, note that $x = 1/(\varphi(x) + n)$ for $n \in \mathbb{N}$. It follows that

$$\begin{aligned} \mu(\varphi^{-1}[a, b]) &= \frac{1}{\log 2} \sum_{n=1}^{\infty} \int_{\frac{1}{n+b}}^{\frac{1}{n+a}} \frac{dx}{1+x} \\ &= \frac{1}{\log 2} \sum_{n=1}^{\infty} \log \frac{1+n+a}{n+a} \frac{n+b}{1+n+b} \\ &= \frac{1}{\log 2} \lim_{n\to\infty} \log \frac{1+n+a}{1+a} \frac{1+b}{1+n+b} \end{aligned}$$

$$= \frac{1}{\log 2} \log \frac{1+b}{1+a}.$$

Thus $\mu(\varphi^{-1}[a,b]) = \mu([a,b])$. Now let $\mathcal{L} = \{A \subset (0,1)\colon \mu(\varphi^{-1}A) = \mu(A)\}$. Then $\mathcal{L}$ is a λ-system (check) and contains all closed intervals. By Dynkin's theorem, $\mu(\varphi^{-1}A) = \mu(A)$ for any Borel set A.

6.2. **Birkhoff's Ergodic Theorem.**

Exercise 6.2.1. As in the proof of Theorem 6.2.1, let $X'_M = X\mathbf{1}_{|X|\le M}$ and $X''_M = X - X'_M$. Following the proof, we know that

$$\frac{1}{n}\sum_{m=0}^{n-1} X'_M(\varphi^m\omega) \to \mathbb{E}[X'_M|\mathcal{I}] \quad \text{a.s.}.$$

Since

$$\left|\frac{1}{n}\sum_{m=0}^{n-1} X'_M(\varphi^m\omega) - \mathbb{E}[X'_M|\mathcal{I}]\right|^p \le (2M)^p,$$

by the bounded convergence theorem, the convergence occurs in L^p as well. To handle X''_M part, observe that,

$$\left\|\frac{1}{n}\sum_{m=0}^{n-1} X''_M(\varphi^m\omega) - \mathbb{E}[X''_M|\mathcal{I}]\right\|_p \le \frac{1}{n}\sum_{m=0}^{n-1} \|X''_M(\varphi^m\omega)\|_p + \|\mathbb{E}[X''_M|\mathcal{I}]\|_p \le 2\left\|X''_M\right\|_p.$$

Let

$$A = \frac{1}{n}\sum_{m=0}^{n-1} X'_M(\varphi^m\omega) - \mathbb{E}[X'_M|\mathcal{I}] \quad \text{and}$$

$$B = \frac{1}{n}\sum_{m=0}^{n-1} X''_M(\varphi^m\omega) - \mathbb{E}[X''_M|\mathcal{I}].$$

It follows that

$$\limsup_{n\to\infty}\left\|\frac{1}{n}\sum_{m=0}^{n-1} X(\varphi^m\omega) - \mathbb{E}[X|\mathcal{I}]\right\|_p \le \limsup_{n\to\infty}(\|A\|_p + \|B\|_p) \le 2\left\|X''_M\right\|_p.$$

Letting $M\to\infty$ and using the dominated convergence theorem, $\mathbb{E}|X''_M|^p \to 0$ and the result follows.

Exercise 6.2.2. (i) Let $h_N = \sup_{m,n\ge N}|g_m - g_n|$. $h_N \le 2\sup_k |g_k|$, so $\mathbb{E}h_N < \infty$. Note that

$$\left|\frac{1}{n}\sum_{m=0}^{n-1} g_m(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right|$$
$$\le \left|\frac{1}{n}\sum_{m=0}^{N-1} g_m(\varphi^m\omega)\right| + \frac{1}{n}\sum_{m=N}^{n-1}|g_m(\varphi^m\omega) - g(\varphi^m\omega)| + \left|\frac{1}{n}\sum_{m=N}^{n-1} g(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right|$$

$$\leq \left|\frac{1}{n}\sum_{m=0}^{N-1} g_m(\varphi^m\omega)\right| + \frac{1}{n}\sum_{m=N}^{n-1} h_N(\varphi^m\omega) + \left|\frac{1}{n}\sum_{m=N}^{n-1} g(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right|.$$

Since $\mathbb{E}\sup_k |g_k| < \infty$, $g_m < \infty$ a.s. and thus the first term $\to 0$ a.s. as $n \to \infty$. Also, since h_N and g are integrable, the second term $\to \mathbb{E}[h_N|\mathcal{I}]$ and the third term $\to 0$ a.s. Thus

$$\limsup_{n\to\infty} \left|\frac{1}{n}\sum_{m=0}^{n-1} g_m(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right| \leq \mathbb{E}[h_N|\mathcal{I}].$$

Now letting $N \to \infty$ and using the dominated convergence, the result follows.

(ii) Note that

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{n}\sum_{m=0}^{n-1} g_m(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right| \\
&\leq \mathbb{E}\left|\frac{1}{n}\sum_{m=0}^{n-1} g_m(\varphi^m\omega) - \frac{1}{n}\sum_{m=0}^{n-1} g(\varphi^m\omega)\right| + \mathbb{E}\left|\frac{1}{n}\sum_{m=0}^{n-1} g(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right| \\
&\leq \frac{1}{n}\sum_{m=0}^{n-1} \mathbb{E}\left|g_m(\varphi^m\omega) - g(\varphi^m\omega)\right| + \mathbb{E}\left|\frac{1}{n}\sum_{m=0}^{n-1} g(\varphi^m\omega) - \mathbb{E}[g|\mathcal{I}]\right|
\end{aligned}$$

Since $g_m \to g$ in L^1, the first term $\to 0$ as $n \to \infty$. The second term $\to 0$ from Theorem 6.2.1.

Exercise 6.2.3. Let $X' = X - \alpha$, $X'_j(\omega) = X'(\varphi^j\omega)$, $S'_k = X'_0 + \cdots + X'_{k-1}$, and $M'_k = \max(0, S'_1, \cdots, S'_k)$. By Lemma 6.2.2, $\mathbb{E}[X'; M'_k > 0] \geq 0$. That is, $\mathbb{E}[X; M'_k > 0] \geq \alpha\mathbb{P}(M'_k > 0)$. Note that $M'_k > 0$ if and only if $D_k > \alpha$. It follows that

$$\mathbb{E}|X| \geq \mathbb{E}[X; M'_k > 0] \geq \alpha\mathbb{P}(M'_k > 0) = \alpha\mathbb{P}(D_k > \alpha).$$

6.3. **Recurrence.**

Exercise 6.3.1. Note that

$$\begin{aligned}
R_n(\omega) &= \mathbf{1}_{\{S_2 - S_1 \neq 0, \cdots, S_n - S_1 \neq 0\}}(\omega) + \cdots + \mathbf{1}_{\{S_n - S_{n-1} \neq 0\}}(\omega) + 1 \\
&= \mathbf{1}_{\{S_1 \neq 0, \cdots, S_{n-1} \neq 0\}}(\varphi\omega) + \cdots + \mathbf{1}_{\{S_1 \neq 0\}}(\varphi^{n-1}\omega) + 1.
\end{aligned}$$

Taking the expectation, we get

$$\begin{aligned}
\mathbb{E}R_n &= \mathbb{P}(S_1 \neq 0, \cdots, S_{n-1} \neq 0) + \cdots + \mathbb{P}(S_1 \neq 0) + 1 \\
&= \sum_{m=1}^{n} g_{m-1}.
\end{aligned}$$

Exercise 6.3.2. Note that, since $\mathbb{P}(X_i > 1) = 0$,

$$\left\{1, \cdots, \max_{m\leq n} S_m\right\} \leq \{S_1, \cdots, S_n\} \leq \left\{\min_{m\leq n} S_m, \cdots, \max_{m\leq n} S_m\right\},$$

and thus

$$\max_{m\le n} S_m \le R_n \le \max_{m\le n} S_m - \min_{m\le n} S_m$$

Since our sequence is ergodic, $S_n/n \to \mathbb{E}X_i > 0$ by the ergodic theorem. This implies that $S_n \to \infty$ and that $\min_m S_m > -\infty$ a.s. Thus,

$$\limsup_{n\to\infty} \max_{m\le n} \frac{S_m}{n} \le \lim_{n\to\infty} \frac{R_n}{n} \le \liminf_{n\to\infty} \max_{m\le n} \frac{S_m}{n},$$

and by Theorem 6.3.1,

$$\mathbb{P}(A) = \lim_{n\to\infty} \max_{m\le n} \frac{S_m}{n}.$$

To evaluate the right-hand-side, note that, for some K,

$$\begin{aligned}
\lim_{n\to\infty} \frac{S_n}{n} &\le \liminf_{n\to\infty} \max_{m\le n} \frac{S_m}{n} \\
&\le \limsup_{n\to\infty} \max_{m\le n} \frac{S_m}{n} \\
&\le \limsup_{n\to\infty} \max_{K\le m\le n} \frac{S_m}{n} \\
&\le \max_{K\le m} \frac{S_m}{m}.
\end{aligned}$$

Since K is arbitrary, by letting $K\to\infty$, $\lim_{n\to\infty}\max_{m\le n} S_m/n = \mathbb{E}X_i$. This completes the proof.

Exercise 6.3.3. Note that

$$\begin{aligned}
\mathbb{E}\left[\sum_{1\le m\le T_1} \mathbf{1}_{\{X_m\in B\}} \middle| X_0 \in A\right] &= \sum_{1\le m} \mathbb{E}\left[\mathbf{1}_{\{X_m\in B\}}\mathbf{1}_{\{T_1\ge m\}} | X_0\in A\right] \\
&= \sum_{1\le m} \mathbb{P}(X_m \in B, T_1 \ge m | X_0 \in A) \\
&= \mathbb{P}(X_0\in A)^{-1} \sum_{1\le m} \mathbb{P}(X_m\in B, T_1\ge m, X_0\in A).
\end{aligned}$$

Let $C_m = \{X_{-m}\in A, X_{-m+1}\notin A, \cdots, X_{-1}\notin A, X_0\in B\}$. Observe that

$$\begin{aligned}
\left(\bigcup_{m=1}^K C_m\right)^c &= \{X_j \in A \text{ for some } -K\le j\le -1, X_0\in B\}^c \\
&= \{X_j\notin A \text{ for all } -K\le j\le -1\}\cup\{X_0\notin B\}.
\end{aligned}$$

As $K\to\infty$, from $\mathbb{P}(X_n\in A \text{ at least once}) = 1$, it follows that $\mathbb{P}(X_j\notin A \text{ for all } -K\le j\le -1) = \mathbb{P}(X_j\notin A \text{ for all } 1\le j\le K)\to 0$. Thus $\mathbb{P}(\bigcup C_m) = \mathbb{P}(X_0\in B)$. Now the first display becomes

$$= \mathbb{P}(X_0\in A)^{-1}\sum_{1\le m}\mathbb{P}(C_m) = \frac{\mathbb{P}(X_0\in B)}{\mathbb{P}(X_0\in A)}.$$

Exercise 6.3.4. From Theorem 6.3.3,

$$\begin{aligned}\frac{\bar{\mathbb{P}}(T_1 \geq n)}{\bar{\mathbb{E}}T_1} &= \mathbb{P}(T_1 \geq n | X_0 = 1)\mathbb{P}(X_0 = 1) \\ &= \mathbb{P}(T_1 \geq n, X_0 = 1).\end{aligned}$$

Note that

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

which completes the proof.

6.5. **Applications.**

Exercise 6.5.1. content

Exercise 6.5.2. content

Exercise 6.5.3. content

Exercise 6.5.4. content

Exercise 6.5.5. content

7. Brownian Motion

7.1. **Definition and Construction.**

Exercise 7.1.1. content

Exercise 7.1.2. Let $X = B_1$, $Y = B_2 - B_1$, and $Z = B_3 - B_2$. Then X, Y, Z are iid normal random variables with mean 0 and variance 1. Note that

$$\begin{aligned}B_1^2 B_2 B_3 &= X^2(X+Y)(X+Y+Z) \\ &= X^4 + 2X^3Y + X^2Y^2 + X^3Z + X^2YZ.\end{aligned}$$

Taking the expectation, it follows that

$$\mathbb{E}B_1^2 B_2 B_3 = \mathbb{E}X^4 + \mathbb{E}X^2Y^2 = 4.$$

7.2. **Markov Property, Blumenthal's 0-1 Law.**

7.3. **Stopping Times, Strong Markov Property.**

7.4. **Path Properties.**

7.5. **Martingales.**

7.6. **Itô's Formula.**

8. Applications to Random Walk

8.1. **Donsker's Theorem.**

8.4. **Empirical Distributions, Brownian Bridge.**

8.5. **Laws of Iterated Logarithm.**

9. Multidimensional Brownian Motion

9.1. **Martingales.**

9.5. **Dirichlet Problem.** end of document

Email address: hoil.lee@kaist.ac.kr

Email address: swj777888@snu.ac.kr