Chapter 7

Exercises

1. Let $\{X_n\}_{n\in\mathbb{N}} \subset \mathcal{L}^p$ for $p \in [1, \infty)$ be a sequence of non-negative functions. Show that

$$\|\sum_{n=1}^{\infty} X_n\|_p \leq \sum_{n=1}^{\infty} \|X_n\|_p.$$

Answer: Let $S_n = \sum_{n=1}^{N} X_n$. Since $X_n \geq 0$ for all n, $0 \leq S_1 \leq S_2 \leq \cdots$. Given that $|S_N|^p \leq 2^p \sum_{n=1}^{N} |X_n|^p$. we have $\int_\Omega |S_N|^p dP \leq 2^p \sum_{n=1}^{N} \int_\Omega |X_n|^p dP < \infty$. Consequently, $S_N \in \mathcal{L}^p$. By Minkowski's inequality

$$\|S_N\|_p \leq \sum_{n=1}^{N} \|X_n\|_p \leq \sum_{n=1}^{\infty} \|X_n\|_p, \tag{7.1}$$

which implies $\|S_N\|_p^p \leq (\sum_{n=1}^{\infty} \|X_n\|_p)^p$. By Beppo-Levi's Theorem

$$\begin{aligned} \sup_{n\in\mathbb{N}} \|S_N\|_p^p = \sup_{n\in\mathbb{N}} \int_\Omega S_N^p dP = \int_\Omega \sup_{n\in\mathbb{N}} S_N^p dP = \int_\Omega \sup_{n\in\mathbb{N}} \left(\sum_{n=1}^{N} X_n \right)^p dP \\ = \int_\Omega \left(\sup_{n\in\mathbb{N}} \sum_{n=1}^{N} X_n \right)^p dP = \|\sum_{n=1}^{\infty} X_n\|_p^p. \end{aligned} \tag{7.2}$$

Hence, by inequality (7.1) and (7.2) we have

$$\sup_{n\in\mathbb{N}} \|S_N\|_p^p = \|\sum_{n=1}^{\infty} X_n\|_p^p \leq \left(\sum_{n=1}^{\infty} \|X_n\|_p \right)^p.$$

Consequently, $\| \sum_{n=1}^{\infty} X_n\|_p \leq \sum_{n=1}^{\infty} \|X_n\|_p$.

2. Show that if $\sum_{n\in\mathbb{N}} x_n$ converges absolutely, then it converges.

Answer: Suppose N_1, $N_2 \in \mathbb{N}$, $N_1 < N_2$ and $\sum\limits_{n\in\mathbb{N}} x_n$ converges absolutely. Note that $\sum_{n=1}^{N_2} |x_n| - \sum_{n=1}^{N_1} |x_n| = \sum_{n=N_1+1}^{N_2} |x_n|$. If $N_1 \to \infty$, then $\sum_{n=1}^{N_2} |x_n| - \sum_{n=1}^{N_1} |x_n| \to 0$, as every convergent sequence is Cauchy. Also, since

$$|x_{N_1+1} + x_{N_2+1} + \cdots x_{N_2}| \leq \sum_{n=N_1+1}^{N_2} |x_n|,$$

$$|\sum_{n=1}^{N_2} x_n - \sum_{n=1}^{N_1} x_n| = |x_{N_1+1} + x_{N_2+1} + \cdots x_{N_2}| \leq \sum_{n=N_1+1}^{N_2} |x_n| \to 0.$$

Since $\mathbb{R}$ is complete, $\lim\limits_{N\to\infty} \sum_{n=1}^{N} x_n$ converges.

3. Prove Theorem 7.9.

   Answer: Let's first prove that $1 \implies 2$. If $\{X_n\}_{n\in\mathbb{N}}$ is uniformly integrable, then X_n is clearly integrable for all n, and consequently $E|X_n| < \infty$ for all n. By the proof of Theorem 7.3, $X_n \xrightarrow{p} X$ implies that there exists a subsequence $X_{n(j)} \xrightarrow{as} X$ as $j \to \infty$. By Fatou's Lemma

$$E|X| = E\left(\liminf_{j\to\infty} |X_{n(j)}|\right) \leq \liminf_{j\to\infty} E|X_{n(j)}| \leq \sup_n E|X_n| < \infty,$$

   where the last inequality follows from Theorem 7.6. Now, for $\epsilon > 0$ let $A_n = \{|X_n - X| > \epsilon\}$

$$\begin{aligned} E|X_n - X| &= E(|X_n - X| I_{|X_n - X| < \epsilon} + |X_n - X| I_{|X_n - X| \geq \epsilon}) \\ &\leq \epsilon + E(|X_n| I_{A_n}) + E(|X| I_{A_n}). \end{aligned}$$

   $P(A_n) \to 0$ as $n \to \infty$, hence $E(|X_n| I_{A_n}) \to 0$ by Theorem 7.6. Similarly, $E(|X| I_{A_n}) \to 0$, which gives the result.

   That $2 \implies 3$ follows from the fact that by the triangle inequality $E|X_n| - E|X| \leq E|X_n - X|$ and $-(E|X_n| - E|X|) \geq -E|X_n - X|$, hence $|E|X_n| - E|X|| \leq E|X_n - X|$.

   Now we prove that $3 \implies 1$. Note that $E(|X_n| I_{|X_n| \geq a}) = E|X_n| - E(|X_n| I_{|X_n| < a}) = E|X_n| - E(|X_n| I_{-a < X_n < a}) = E|X_n| - E(u(X_n))$, if we let $u(x) = |x| I_{(-a,a)}(x)$. By Theorem 7.15 $X_n \xrightarrow{p} X \implies X_n \xrightarrow{d} X$. But since $u(x)$ is bounded and continuous, we have that using Definition 7.5 and by Theorem 7.11,

$$E(u(X_n)) \to E(u(X))$$

if a and $-a$ are points of continuity of the distribution F_X of X. Hence, for a a point of continuity of F_X, we have

$$E(|X_n|I_{|X_n|\geq a}) \to E(|X|I_{|X|\geq a}).$$

For any $\epsilon > 0$ there exists a point of continuity b of F_X such that $E(|X|I_{|X|\geq b}) < \epsilon$. Using this b and $E(|X_n|I_{|X_n|\geq b}) = E|X_n| - E(|X_n|I_{-b<X_n<b})$ there exists N such that for all $n > N$,

$$E(|X_n|I_{|X_n|\geq b}) < 2\epsilon.$$

Also, there exists $c > 0$ such that $E(|X_k|I_{|X_k|\geq c}) < 2\epsilon$ for all $k < N$ since there are only finitely many terms involved. Hence, if $a > \max\{b, c\}$ we have uniform integrability.

4. Let $\{g_n\}_{n=1,2,\cdots}$ be a sequence of real valued functions that converge uniformly to g on an open set S, containing x, and g is continuous at x. Show that if $\{X_n\}_{n=1,2,\cdots}$ is a sequence of random variables taking values in S such that $X_n \xrightarrow{p} X$, then

$$g_n(X_n) \xrightarrow{p} g(X).$$

Note: Recall that a sequence of real valued functions $\{g_n\}_{n=1,2,\cdots}$ converges uniformly to g on a set S if, for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ (depending only on ϵ) such that for all $n > N_\epsilon$, $|g_n(x) - g(x)| < \epsilon$ for every $x \in S$.

Answer: Let $\epsilon, \delta > 0$ and define the following subsets of the sample space: $S_1^n = \{\omega : |g_n(X_n) - g(X)| < \epsilon\}$, $S_2^n = \{\omega : |g_n(X_n) - g(X_n)| < \epsilon/2\}$, $S_3^n = \{\omega : |g(X_n) - g(X)| < \epsilon/2\}$, $S_4^n = \{\omega : X_n \in S\}$. By the triangle inequality, $S_1^n \supseteq S_2^n \cap S_3^n$. By continuity of g at X and openness of S, there exists γ_ϵ such that whenever $|X_n - X| < \gamma_\epsilon$, $|g(X_n) - g(X)| < \epsilon/2$ and $X_n \in S$. Letting, $S_5^n = \{\omega : |X_n - X| < \gamma_\epsilon\}$, we see that $S_5^n \subseteq S_3^n \cap S_4^n$. Since $X_n \xrightarrow{p} X$ and uniform convergence of g_n, there exists $N_{\delta,\epsilon}$ such that whenever $n > N_{\delta,\epsilon}$, $|g_n(X) - g(X)| < \epsilon/2$ for all $X \in S$ and $P(S_5^n) > 1 - \delta$. Thus, $n > N_{\delta,\epsilon}$ implies $S_4^n \subseteq S_2^n$. Consequently, $n > N_{\delta,\epsilon}$ implies $S_1^n \supseteq S_2^n \cap S_3^n \supseteq S_4^n \cap S_3^n \supseteq S_5^n$. Thus, $P(S_1^n) \geq P(S_5^n) > 1 - \delta$.

5. Show that $X_n \xrightarrow{as} X$ is equivalent to $P\left(\{\omega : \sup_{j\geq n} |X_j - X| \geq \epsilon\}\right) \to 0$ for all $\epsilon > 0$ as $n \to \infty$.

Answer: For any $\epsilon > 0$ and $k \in \mathbb{N}$ let $A_k(\epsilon) = \{\omega : |X_k(\omega) - X(\omega)| > \epsilon\}$. If for all $n \in \mathbb{N}$ we have that $P\left(\cup_{k>n} A_k(\epsilon)\right) > 0$ then it must be that $X_n \overset{as}{\nrightarrow} X$. Consequently,

$$\begin{aligned} X_n \overset{as}{\to} X \quad &\Leftrightarrow \quad \lim_{n\to\infty} P\left(\cup_{n<k} A_k(\epsilon)\right) = 0 \\ &\Leftrightarrow \quad P\left(\{\omega : \sup_{j\geq n} |X_j - X| > \epsilon\}\right) \to 0 \text{ as } n \to \infty. \end{aligned}$$

6. Prove item 1 of Remark 7.1.

   Answer: For $\epsilon > 0$ we have that

$$\{\omega : |X_n + Y_n - X - Y| > \epsilon\} \subseteq \{\omega : |X_n - X| > \epsilon/2\} \cup \{\omega : |Y_n - Y| > \epsilon/2\}$$

   The probability of the events on the union on right-hand side go to zero as $n \to \infty$. By monotonicity of probability measures we have the results.

7. Let $n \in \mathbb{N}$ and $h_n > 0$ such that $h_n \to 0$ as $n \to \infty$. Show that if $\sum_{n=1}^{\infty} P(\{\omega : |X_n - X| \geq h_n\}) < \infty$ then $X_n \overset{p}{\to} X$.

   Answer: From question 5,

$$X_n \overset{as}{\to} X \quad \Leftrightarrow \quad \lim_{n\to\infty} P\left(\cup_{n<k} A_k(h_n)\right) = 0.$$

   But $P\left(\cup_{n<k} A_k(h_n)\right) \leq \sum_{k\geq n} P(A_k(\epsilon))$ and if $\sum_{n=1}^{\infty} P(\{\omega : |X_n - X| \geq h_n\}) < \infty$ then it must be that $\lim_{n\to\infty} \sum_{k\geq n} P(A_k(\epsilon)) = 0$. Since convergence almost surely implies convergence in probability, the proof is complete.

8. Show that if $Y_n \overset{d}{\to} Y$ then $Y_n = O_p(1)$.

   Answer: Without loss of generality let $a > 0$. Provided that a and $-a$ are continuity points of F_Y, we can write that,$P(|Y_n| > a) \to P(|Y| > a)$ as $n \to \infty$. Hence, for every $\epsilon > 0$ there exists N_ϵ such that,

$$|P(|Y_n| > a) - P(|Y| > a)| < \epsilon \text{ for all } n \geq N_\epsilon$$

   or

$$P(|Y| > a) - \epsilon < P(|Y_n| > a) < P(|Y| > a) + \epsilon.$$

We can choose a such that $P(|Y| > a) < \delta$ for any $\delta > 0$. Thus, $P(|Y_n| > a) < \delta + \epsilon$ for all $n \geq N_\epsilon$.

9. Let $g : S \subseteq \mathbb{R}$ be continuous on S, and X_t and X_s be random variables defined on $(\Omega, \mathcal{F}, P)$ taking values in S. Show that: a) if X_t is independent of X_s, then $g(X_t)$ is independent of $g(X_s)$; b) if X_t and X_s are identically distributed, then $g(X_t)$ and $g(X_s)$ are identically distributed.

   Answer: Let $Y_t = g(X_t)$ and $Y_s = g(X_s)$. g continuous assures that both Y_t and Y_s are random variables.

   a) $F_{Y_t,Y_s}(a, b) = P(S = \{\omega : Y_t \leq a \text{ and } Y_s \leq b\})$. Let $S_t = \{X_t(\omega) : Y_t(\omega) \leq a\}, S_s = \{X_s(\omega) : Y_s(\omega) \leq b\}$. Since, $S = S_t \cap S_s$ and by independence $P(S) = P(S_t)P(S_s)$ which implies $F_{Y_t,Y_s}(a, b) = F_{Y_t}(a)F_{Y_s}(b)$.

   b) $F_{Y_t}(a) = P(S_t) = P(\{X_s(\omega) : Y_s(\omega) \leq a\}) = F_{Y_s}(a)$.

10. Let $\{X_n\}$ be a sequence of independent random variables that converges in probability to a limit X. Show that X is almost surely a constant.

   Answer: Recall that if X is almost surely a constant, say c, $P(\{\omega : X(\omega) \neq c\}) = 0$. Then, the distribution function F associated with X is given by

$$F(x) = \begin{cases} 0, & \text{if } x < c \\ 1, & \text{if } x \geq c \end{cases} .$$

   If X is not a constant, there exists a c and $0 < \epsilon < 1/2$ such that $P(X < c) > 2\epsilon$ and $P(X \leq c + \epsilon) < 1 - 2\epsilon$ or $P(X > c + \epsilon) > 2\epsilon$. Since $X_n \xrightarrow{p} X$ than $X_n \xrightarrow{d} X$. Consequently, for n sufficiently large and c a point of continuity of F we have

$$F(c) - \epsilon < F_n(c) < F(c) + \epsilon$$

   which implies that $\epsilon < F_n(c)$. Also, $1 - F_n(c + \epsilon) > 1 - F(c + \epsilon) - \epsilon$ which implies $P(X_n > c + \epsilon) > P(X > c + \epsilon) - \epsilon > \epsilon$. Since $X_n \xrightarrow{p} X$, for n sufficiently large

$P(\{\omega : |X_r - X_s| > \epsilon\}) < \epsilon^3$. Since $\{\omega : |X_r - X_s| > \epsilon\} = \{\omega : X_r - X_s > \epsilon\} \cup \{\omega : X_r - X_s < -\epsilon\}$ we note that if $X_r < c$ and $X_s > c + \epsilon$ then $X_r - X_s > \epsilon$ is equivalent to $X_r - X_s < -\epsilon$. Consequently,

$$P(\{\omega : |X_r - X_s| > \epsilon\}) \leq P(\{\omega : X_r < c \text{ and } X_s > c + \epsilon\}).$$

But since X_r and X_s are independent $P(\{\omega : X_r < c \text{ and } X_s > c + \epsilon\}) = P(\{\omega : X_r < c\})P(\{\omega : X_s > c + \epsilon\}) > \epsilon^2$. Hence,

$$\epsilon^3 > P(\{\omega : |X_r - X_s| > \epsilon\}) > \epsilon^2,$$

a contradiction.

11. Suppose $\frac{X_n - \mu}{\sigma_n} \xrightarrow{d} Z$ where the non-random sequence $\sigma_n \to 0$ as $n \to \infty$, and g is a function which is differentiable at μ. Then, show that $\frac{g(X_n) - g(\mu)}{g^{(1)}(\mu)\sigma_n} \xrightarrow{d} Z$.

    Answer: From question 2, if $Z_n \xrightarrow{d} Z$ then $Z_n = O_p(1)$. Let $Z_n = \frac{X_n - \mu}{\sigma_n}$ and write $X_n = \mu + \sigma_n Z_n = \mu + O_p(\sigma_n)$. By Taylor's Theorem

    $$\frac{1}{\sigma_n} g(X_n) - g(\mu) = g^{(1)}(\mu)\frac{(X_n - \mu)}{\sigma_n} + o_p(1).$$

    Since $\frac{X_n - \mu}{\sigma_n} \xrightarrow{d} Z$, we have the result.

12. Show that if $\{X_n\}_{n \in \mathbb{N}}$ and X are random variables defined on the same probability space and $r > s \geq 1$ and $X_n \xrightarrow{\mathcal{L}_r} X$, then $X_n \xrightarrow{\mathcal{L}_s} X$.

    Answer: For arbitrary W let $Z = |W|^s$, $Y = 1$ and $p = r/s$. Then, by Hölder's Inequality

    $$E|ZY| \leq \|Z\|_p \|Y\|_{p/(p-1)}.$$

    Substituting Z and Y gives $E(|W|^s) \leq E(|W|^{sp})^{1/p} = E(|W|^{s\frac{r}{s}})^{s/r}$. Raising both sides to $1/s$ gives

    $$E(|W|^s)^{1/s} \leq E(|W|^r)^{1/r}.$$

    Setting $W = X_n - X$ and taking limits as $n \to \infty$ gives the result.

13. Let $\mathcal{L}_P^\infty := \{X : \Omega \to \mathbb{R}$ such that, there exists $C > 0$, with $P(\{\omega : |X(\omega)| \geq C\}) = 0\}$. For this space of random variables define the norm $\|X\|_\infty := \inf\{C > 0 : P(\{\omega : |X(\omega)| \geq C\}) = 0\}$. Establish that this space is complete.

Answer: If $\{X_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}_P^\infty$, we can set

$$A_{k,l} := \{|X_k| > \|X_k\|_\infty\} \cup \{|X_k - X_l| > \|X_k - X_l\|_\infty\}, A := \bigcup_{k,l\in\mathbb{N}} A_{k,l}.$$

By definition $P(A_{k,l}) = 0$ and $P(A) = 0$, so that $\|X_n I_A\|_\infty = 0$ for all n. On the set A^c, $\{X_n\}_{n\in\mathbb{N}}$ converges uniformly to a bounded function X, i.e., $XI_{A^c} \in \mathcal{L}_P^\infty$ as well as $\|(X_n - X)I_{A^c}\|_\infty \to 0$.

Chapter 8

Exercises

1. Let U and V be two points in an n-dimensional unit cube, i.e., $[0,1]^n$ and X_n be the Euclidean distance between these two points which are chosen independently and uniformly. Show that $\frac{X_n}{\sqrt{n}} \xrightarrow{p} \frac{1}{\sqrt{6}}$.

   Answer: Let $U' = \begin{pmatrix} U_1 & \cdots & U_n \end{pmatrix}$ and $V' = \begin{pmatrix} V_1 & \cdots & V_n \end{pmatrix}$. Then, $X_n = \left(\sum_{i=1}^n (U_i - V_i)^2\right)^{1/2}$ and we can write

   $$\frac{1}{n}E(X_n^2) = \frac{1}{n}\sum_{i=1}^n E((U_i - V_i)^2) = \int_0^1 \int_0^1 (u-v)^2 du dv = 1/6$$

   where the last equality follows from routine integration. Then, since $E(|(U-V)^2|) = E((U-V)^2) < \infty$, by Kolmogorov's Law of Large Numbers

   $$\frac{1}{n}X_n^2 = \frac{1}{n}\sum_{i=1}^n (U_i - V_i)^2 \xrightarrow{p} 1/6.$$

   Since, $f(x) = x^{1/2}$ is a continuous function $[0,\infty)$, by Slutsky Theorem if $\frac{1}{n}X_n^2 \xrightarrow{p} 1/6$ then $f\left(\frac{1}{n}X_n^2\right) \xrightarrow{p} f(1/6)$. Consequently,

   $$\frac{1}{\sqrt{n}}X_n \xrightarrow{p} 1/\sqrt{6}.$$

2. Show that if $\{X_j\}_{j\in\mathbb{N}}$ be a sequence of random variables with $E(X_j) = 0$ and $\sum_{j=1}^\infty \frac{1}{a_j^p} E(|X_j|^p) < \infty$ for some $p \geq 1$ and a sequence of positive constants $\{a_j\}_{j\in\mathbb{N}}$. Then,

   $$\sum_{j=1}^\infty P(|X_j| > a_j) < \infty \text{ and } \sum_{j=1}^\infty \frac{1}{a_j}|E(X_j I_{\{\omega:|X_j|\leq a_j\}})| < \infty.$$

Furthermore, for any $r \geq p$,

$$\sum_{j=1}^{\infty} \frac{1}{a_j^r} E(|X_j|^r I_{\{\omega:|X_j|\leq a_j\}}) < \infty.$$

Use this result to prove Theorem 8.4 in your class notes with convergence in probability.

Answer: Note that

$$P\left(\{\omega : |X_j| > a_j\}\right) = 1 - P(\{\omega : |X_j| \leq a_j\}) = \int_{\Omega} \left(1 - I_{\{\omega:|X_j|\leq a_j\}}\right) dP.$$

If $\omega \in \{\omega : |X_j| \leq a_j\}$, then $P\left(\{\omega : |X_j| > a_j\}\right) = 0$. If $|X_j| > a_j$, then $|X_j|^p > a_j^p$ and $|X_j|^p/a_j^p > 1$. Hence,

$$P\left(\{\omega : |X_j| > a_j\}\right) < \int_{\Omega} |X_j|^p/a_j^p dP = \frac{1}{a_j^p} E\left(|X_j|^p\right)$$

and

$$\sum_{j=1}^{\infty} P\left(\{\omega : |X_j| > a_j\}\right) < \sum_{j=1}^{\infty} \frac{1}{a_j^p} E\left(|X_j|^p\right) < \infty.$$

Now,

$$\begin{aligned}
\frac{1}{a_j}|E(X_j I_{\{\omega:|X_j|\leq a_j\}})| &= \frac{1}{a_j}|E(X_j) - E(X_j I_{\{\omega:|X_j|\leq a_j\}})|, \text{ since } E(X_j) = 0. \\
&\leq \frac{1}{a_j} E\left(|X_t|(1 - I_{\{\omega:|X_j|\leq a_j\}}\right) \\
&\leq \frac{1}{a_j^p} E\left(|X_j|^p(1 - I_{\{\omega:|X_j|\leq a_j\}}\right) \text{ since } \tfrac{|X_j|^p}{a_j^p} \geq \tfrac{|X_j|}{a_j} \text{ if } p \geq 1. \\
&\leq \frac{1}{a_j^p} E\left(|X_j|^p\right).
\end{aligned}$$

Hence,

$$\sum_{j=1}^{\infty} \frac{1}{a_j}|E(X_j I_{\{\omega:|X_j|\leq a_j\}})| < \sum_{j=1}^{\infty} \frac{1}{a_j^p} E\left(|X_j|^p\right) < \infty.$$

Lastly, if $|X_j| \leq a_j$ we have that $\frac{1}{a_j}|X_j| \leq 1$. Then, for $r \geq p \geq 1$

$$\frac{1}{a_j^r}|X_j|^r I_{\{\omega:|X_j|\leq a_j\}} \leq \frac{1}{a_j^p}|X_j|^p I_{\{\omega:|X_j|\leq a_j\}} \leq \frac{1}{a_j}|X_j| I_{\{\omega:|X_j|\leq a_j\}}$$

and

$$E\left(\frac{1}{a_j^r}|X_j|^r I_{\{\omega:|X_j|\leq a_j\}}\right) \leq E\left(\frac{1}{a_j^p}|X_j|^p I_{\{\omega:|X_j|\leq a_j\}}\right).$$

Hence,

$$\sum_{j=1}^{\infty} E\left(\frac{1}{a_j^r}|X_j|^r I_{\{\omega:|X_j|\leq a_j\}}\right) < \infty.$$

In Theorem 8.4, the sequence of random variables $\{X_j\}_{j\in\mathbb{N}}$ is independent and has expectation μ_j. Hence, if $W_j := X_j - \mu_j$, we have $E(W_j) = 0$. Furthermore, in Theorem 8.4 it is assumed that for some $\delta > 0$

$$\sum_{j=1}^{\infty} \frac{E(|W_j|^{1+\delta})}{j^{1+\delta}} < \infty.$$

Now, note that for any $n \in \mathbb{N}$ we have $\sum_{j=1}^{n} \frac{E(|W_j|^{1+\delta})}{n^{1+\delta}} \leq \sum_{j=1}^{n} \frac{E(|W_j|^{1+\delta})}{j^{1+\delta}}$ and

$$\lim_{n\to\infty} \sum_{j=1}^{n} \frac{E(|W_j|^{1+\delta})}{n^{1+\delta}} \leq \lim_{n\to\infty} \sum_{j=1}^{n} \frac{E(|W_j|^{1+\delta})}{j^{1+\delta}} < \infty.$$

Now, in the first part of this answer, take $a_j = n$ for all j and for any $r > 1+\delta$. Then, we have

$$\sum_{j=1}^{\infty} P(|W_j| > n) < \infty \text{ and } \sum_{j=1}^{\infty} \frac{1}{n^r} E(|W_j|^r I_{\{\omega:|W_j|\leq n\}}) < \infty.$$

Hence, taking $r = 2$ the conditions on Theorem 8.2 are met and we have

$$\frac{1}{n}\sum_{j=1}^{n} W_j - \frac{1}{n}\sum_{i=1}^{n} E\left(W_j I_{\{\omega:|W_j|\leq n\}}\right) = \frac{1}{n}\sum_{j=1}^{n}(X_j - \mu_j) - \frac{1}{n}\sum_{i=1}^{n} E\left(W_j I_{\{\omega:|W_j|\leq n\}}\right) = o_p(1).$$

But since $E(W_j) = 0$, we have $E\left(W_j I_{\{\omega:|W_j|\leq n\}}\right) \to 0$ as $n \to \infty$. Thus, $\frac{1}{n}\sum_{j=1}^{n}(X_j - \mu_j) = o_p(1)$.

3. Let $\{X_i\}_{i=2,3,\dots}$ be a sequence of independent random variables such that

$$P(X_i = i) = P(X_i = -i) = \frac{1}{2i\log i},\ P(X_i = 0) = 1 - \frac{1}{i\log i}$$

Show that $\frac{1}{n}\sum_{i=2}^{n} X_i \xrightarrow{p} 0$.

Answer: Let $S_n = \sum_{i=2}^{n} X_i$ and note that $E(X_i) = 0$. Hence, by independence

$$E(S_n^2) = \sum_{i=2}^{n} E(X_i^2) = \sum_{i=2}^{n} \frac{i}{\log i} \leq \frac{n^2}{\log n}.$$

Hence, $V(S_n/n) = \frac{1}{n^2}V(S_n) = \frac{1}{n^2}E(S_n^2) \leq \frac{1}{n^2}\frac{n^2}{\log n} = \frac{1}{\log n} \to 0$ as $n \to \infty$. Consequently, $\frac{1}{n}S_n \xrightarrow{p} 0$ by Chebyshev's inequality.

Chapter 9

Exercises

1. Assess the veracity of the following statement: "Since knowledge of X implies knowledge of $f(X)$, conditioning on X is the same as conditioning on $f(X)$. Hence, $E(Y|f(X)) = E(Y|X)$." Explain using mathematical arguments.

   Answer: The statement is false. Recall that conditioning on a random variable X means conditioning on the sub-σ-algebra generated by X, i.e., $X^{-1}(\mathcal{B})$. Hence, conditioning on $f(X)$ means conditioning on the sub-σ-algebra generated by $f(X)$, i.e., $X^{-1}(f^{-1}(\mathcal{B}))$ which is generally different from $X^{-1}(\mathcal{B})$. Take, for example, the following random vector: $(Y, X) : \Omega \to \mathbb{R}^2$ with $(Y(\omega), X(\omega)) = (1, -1)$ if $\omega \in E_1$ and $(Y(\omega), X(\omega)) = (2, -1)$ if $\omega \in E_2$, $(Y(\omega), X(\omega)) = (1, 1)$ if $\omega \in E_3$ and $(Y(\omega), X(\omega)) = (2, 1)$ if $\omega \in E_4$, with $P(E_j) = 1/6$ for $j = 1, 2$, $P(E_3) = 3/6$, $P(E_4) = 1/6$ and $\Omega = \cup_{j=1}^{4} E_j$ and $E_i \cup E_j = \emptyset$ for $i \neq j$. Now, let $f(X) = X^2$. Then,

   $$E(Y|X) = \begin{cases} 1.5 & \text{if } X = -1 \\ 5/4 & \text{if } X = 1 \end{cases} \quad \text{and } E(Y|X^2) = 8/6.$$

2. Let X and Y be independent random variables defined in the same probability space. Show that if $E(|Y|) < \infty$ then

   $$P\left(E(Y|X) = E(Y)\right) = 1.$$

   Answer: Let $\mathcal{F}_X$ be the σ-algebra generated by X. Let $E \in \mathcal{F}_X$ and note that there exists B such that $E = \{\omega : X(\omega) \in B\}$.

   $$\int_A Y dP = \int_\Omega Y I_A dP = \int_\Omega Y I_{X \in B} dP = E(Y I_{X \in B}) = E(Y)E(I_{X \in B})$$

where the last equality follows by independence. Now,

$$E(Y)E(I_{X\in B}) = E(Y)\int_{\Omega} I_{X\in B}dP = E(Y)\int_{\Omega} I_A dP = \int_A E(Y)dP.$$

Consequently, since A is arbitrary in $\mathcal{F}_X$

$$\int_A Y dP = \int_A E(Y)dP \text{ or } \int_A (Y - E(Y))dP = 0$$

By definition of conditional expectation we have that $E(Y|X) = E(Y)$ since A is arbitrary in $\mathcal{F}_X$.

3. Let $(\Omega, \mathcal{F}, P)$ be a probability space. The set of random variables $X : \Omega \to \mathbb{R}$ such that $\int_\Omega X^2 dP < \infty$ is denoted by $L^2(\Omega, \mathcal{F}, P)$. On this set $\|X\| = \left(\int_\Omega X^2 dP\right)^{1/2}$ is a norm and $< X, Y >= \int_\Omega XY dP$ is an inner product. If $\mathcal{G}$ is a σ-algebra and $\mathcal{G} \subset \mathcal{F}$, the conditional expectation of X with respect to $\mathcal{G}$, denoted by $E(X|\mathcal{G})$ is the orthogonal projection of X onto the closed subspace $L^2(\Omega, \mathcal{G}, P)$ of $L^2(\Omega, \mathcal{F}, P)$. Prove the following results:

   (a) For $X, Y \in L^2(\Omega, \mathcal{F}, P)$ we have $< E(X|\mathcal{G}), Y >=< E(Y|\mathcal{G})), X >=< E(X|\mathcal{G}), E(Y|\mathcal{G}) >$.

   (b) If $X = Y$ almost everywhere then $E(X|\mathcal{G}) = E(Y|\mathcal{G})$ almost everywhere.

   (c) For $X \in L^2(\Omega, \mathcal{G}, P)$ we have $E(X|\mathcal{G}) = X$.

   (d) If $\mathcal{H} \subset \mathcal{G}$ is a σ-algebra, then $E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$.

   (e) If $Y \in L^2(\Omega, \mathcal{G}, P)$ and there exists a constant $C > 0$ such that $P(|Y| \geq C) = 0$, we have that $E(YX|\mathcal{G}) = YE(X|\mathcal{G})$.

   (f) If $\{Y_n\}_{n\in\mathbb{N}}, X \in L^2(\Omega, \mathcal{F}, P)$ and $\|Y_n - X\| \to 0$ as $n \to \infty$, then $E(Y_n|\mathcal{G}) \xrightarrow{p} E(X|\mathcal{G})$ as $n \to \infty$.

Answer: (a) By definition of conditional expectation, for all measurable $s \in L^2(\Omega, \mathcal{G}, P)$,

$$E\left([X - E(X|\mathcal{G})]s\right) = 0 \iff E(Xs) = E\left(E(X|\mathcal{G})s\right). \tag{9.1}$$

Since $E(Y|\mathcal{G}) \in L^2(\Omega, \mathcal{G}, P)$, we have $E(XE(Y|\mathcal{G})) = E\left(E(X|\mathcal{G})E(Y|\mathcal{G})\right)$. But by definition of the inner product the last equality is $< E(Y|\mathcal{G})), X >=< E(X|\mathcal{G}), E(Y|\mathcal{G}) >$. Similarly, changing X for Y in equation (9.1) we obtain $E(Ys) = E\left(E(Y|\mathcal{G})s\right)$. Letting, $s = E(X|\mathcal{G})$ we get $E(YE(X|\mathcal{G})) = E\left(E(Y|\mathcal{G})E(X|\mathcal{G})\right)$ and $E(YE(X|\mathcal{G})) = E(XE(Y|\mathcal{G}))$, which is equivalent to $< E(X|\mathcal{G}), Y >=< E(Y|\mathcal{G})), X >$.

(b) Let $y = E(Y|\mathcal{G})$ and $x = E(X|\mathcal{G})$. Then,

$$\begin{aligned}(y-x)^2 &= (y-Y+Y-x)(y-x) = (y-Y)(y-x) + (Y-x)(y-x) \\ &= (y-Y)(y-x) + (Y-X)(y-x) + (X-x)(y-x)\end{aligned}$$

But from item 1, $E(y-x)^2 := \|y-x\|^2 = E(Y-X)(y-x) \leq E|(Y-X)(y-x)| \leq \|Y-X\|\|y-x\|$, which gives $\|y-x\| \leq \|Y-X\|$. Lastly, if $X = Y$ almost everywhere, then $\|Y-X\| = 0$ and $x = y$ almost everywhere.

(c) Since $X \in L^2(\Omega, \mathcal{G}, P)$, it follows from the projection theorem that $E(X|\mathcal{G}) = X$.

(d) From item (a), we have $< E(E(X|\mathcal{G})|\mathcal{H}), Y >=< E(X|\mathcal{G}), E(Y|\mathcal{H}) >=< X, E(E(Y|\mathcal{H})|\mathcal{G}) >$. Since $E(Y|\mathcal{H}) \in L^2(\Omega, \mathcal{G}, P)$, we have that by item (c) $< X, E(E(Y|\mathcal{H})|\mathcal{G}) >=< X, E(Y|\mathcal{H}) >=< E(X|\mathcal{H}), Y >$. Hence, $E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$ almost everywhere.

(e) Since $L^2(\Omega, \mathcal{G}, P)$ is a closed linear subspace of $L^2(\Omega, \mathcal{F}, P)$ and $E(\cdot|\mathcal{G})$ is a linear projector, any $X \in L^2(\Omega, \mathcal{F}, P)$ can be written as

$$X = E(X|\mathcal{G}) + (X - E(X|\mathcal{G})) \tag{9.2}$$

where $(X - E(X|\mathcal{G}))$ is orthogonal to any element of $L^2(\Omega, \mathcal{G}, P)$. Hence, (9.2) gives

$$XY = E(X|\mathcal{G})Y + (X - E(X|\mathcal{G}))Y. \tag{9.3}$$

Now, note that for any $s \in L^2(\Omega, \mathcal{G}, P)$ and $Y \in L^2(\Omega, \mathcal{G}, P)$ bounded almost everywhere, as assumed in the question, we have $sY \in L^2(\Omega, \mathcal{G}, P)$. Hence, $E((X - E(X|\mathcal{G}))sY) = 0$ and using (9.3) we have

$$E(sXY) = E(sE(X|\mathcal{G})Y) \iff E([XY - E(X|\mathcal{G})Y]s) = 0,$$

and the conclusion that $E(XY|\mathcal{G}) = E(X|\mathcal{G})Y$.

(f) From item (b)

$$\|E(Y_n|\mathcal{G}) - E(Z|\mathcal{G})\| \leq \|Y_n - Z\|.$$

Taking limits on both sides as $n \to \infty$ we obtain $\|E(Y_n|\mathcal{G}) - E(Z|\mathcal{G})\| \to 0$, since $\|Y_n - Z\| \to 0$ by assumption. That is, $E(Y_n|\mathcal{G})$ converges in quadratic mean to $E(Z|\mathcal{G})$. But by Chebyshev's inequality, convergence in quadratic mean implies convergence in probability. Hence, $E(Y_n|\mathcal{G}) \xrightarrow{p} E(Z|\mathcal{G})$.

4. Let $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ be random variables and assume that $E(Y|X) = aX$ where $a \in \mathbb{R}$.

   (a) Show that if $E(X^2) > 0$, $a = E(XY)/E(X^2)$.

   (b) If $\{(Y_i \ X_i)^T\}_{i=1}^n$ is a sequence of independent random vectors with components having the same distribution as $(Y \ X)^T$, show that

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{p} E(X^2) \text{ and } \frac{1}{n}\sum_{i=1}^n Y_i X_i \xrightarrow{p} E(XY).$$

   (c) Let $a_n = \left(\frac{1}{n}\sum_{i=1}^n X_i^2\right)^{-1} \frac{1}{n}\sum_{i=1}^n Y_i X_i$. Does $a_n \xrightarrow{p} a$? Can a_n be defined for all n? Explain.

Answer: (a) Note that $E(Y|X) = \underset{a}{\operatorname{argmin}} \int_\Omega (Y - aX)^2 dP$. Now,

$$\int_\Omega (Y - aX)^2 dP = \int_\Omega Y^2 dP + a^2 \int_\Omega X^2 dP - 2a \int_\Omega XY dP,$$

$$\frac{d}{da}\int_\Omega (Y - aX)^2 dP = 2a \int_\Omega X^2 dP - 2\int_\Omega XY dP \text{ and } \frac{d^2}{da^2}\int_\Omega (Y - aX)^2 dP = 2\int_\Omega X^2 dP > 0.$$

Hence, setting the first derivative equal to zero gives, $E(Y|X) \int_\Omega X^2 dP = \int_\Omega XY dP \iff E(Y|X) = \frac{E(XY)}{E(X^2)}$.

(b) Since $X_i^2 = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \begin{pmatrix} Y_i \\ X_i \end{pmatrix}^T \begin{pmatrix} 0 & 1 \end{pmatrix}^T$ and $X_i Y_i = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \begin{pmatrix} Y_i \\ X_i \end{pmatrix}^T (0 \quad 1$ they are measurable function of $\begin{pmatrix} Y_i \\ X_i \end{pmatrix}$. Hence, $\{X_i^2\}_{i\in\mathbb{N}}$ and $\{X_i Y_i\}_{i\in\mathbb{N}}$ are IID sequences. Since, $E(X_i^2) = E(X^2)$ and $E(X_i Y_i) = E(XY)$ by the law of large numbers for IID random variables

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{p} E(X^2) > 0 \text{ and } \frac{1}{n}\sum_{i=1}^n Y_i X_i \xrightarrow{p} E(XY).$$

(c) To define a_n we need $\frac{1}{n}\sum_{i=1}^n X_i^2 > 0$ which is not assured from the assumptions. What can be said is that $\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{p} E(X^2) > 0$. Hence, a_n exists in probability as $n \to \infty$.

5. Prove the following:

(a) If $Y \in \mathcal{L}(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subset \mathcal{F}$ is a σ-algebra, show that $|E(Y|\mathcal{G})| \leq E(|Y||\mathcal{G})$.

Answer: $|Y| = Y^+ - Y^-$ where Y^+, $Y^- \geq 0$. By linearity of conditional expectation

$$E(|Y||\mathcal{G}) = E(Y^+|\mathcal{G}) + E(Y^-|\mathcal{G})$$

and from Theorem 7.9 $E(Y^+|\mathcal{G}) \geq 0$, $E(Y^-|\mathcal{G}) \geq 0$. Hence,

$$\begin{aligned}|E(Y|\mathcal{G})| &= |E(Y^+|\mathcal{G}) - E(Y^-|\mathcal{G})| \leq |E(Y^+|\mathcal{G})| + |E(Y^-|\mathcal{G})| \\ &= E(Y^+|\mathcal{G}) + E(Y^-|\mathcal{G}) = E(|Y||\mathcal{G})\end{aligned}$$

(b) Let c be a scalar constant and suppose $X = c$ almost surely. Show that $E(X|\mathcal{G}) = c$ almost surely.

Answer: It suffices to show that $\int_\Omega |c - E(X|\mathcal{G})|dP = 0$. Now,

$$\int_\Omega |c - E(X|\mathcal{G})|dP \;=\; \int_{c \geq E(X|\mathcal{G})} (c - E(X|\mathcal{G}))dP + \int_{c < E(X|\mathcal{G})} (E(X|\mathcal{G}) - c)dP.$$

Now, $\int_{c \geq E(X|\mathcal{G})}(c - E(X|\mathcal{G}))dP = \int_\Omega (c - E(X|\mathcal{G}))I_{\{c \geq E(X|\mathcal{G})\}}dP$. Now, since $E(X|\mathcal{G}) \in \mathcal{L}(\Omega, \mathcal{G}, P)$, $I_{\{c \geq E(X|\mathcal{G})\}}$ is $\mathcal{G}$-measurable. Hence, by the definition of conditional expectation

$$\int_{c \geq E(X|\mathcal{G})} (c - E(X|\mathcal{G}))dP = 0.$$

Similarly, $\int_{c < E(X|\mathcal{G})}(E(X|\mathcal{G}) - c)dP = 0$. Hence, $c = E(X|\mathcal{G})$ almost surely.

(c) If $Y \in \mathcal{L}(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subset \mathcal{F}$ is a σ-algebra, show that for $a > 0$

$$P\left(\{\omega : |Y(\omega)| \geq a\}|\mathcal{G}\right) \leq \frac{1}{a}E(|Y(\omega)||\mathcal{G}).$$

What is the definition of $P\left(\{\omega : |Y(\omega)| \geq a\}|\mathcal{G}\right)$? Is this a legitimate probability measure?

Answer: Note that $aI_{\{\omega:|Y(\omega)|\geq a\}} \leq |Y(\omega)|$ and

$$aE\left(I_{\{\omega:|Y(\omega)|\geq a\}}|\mathcal{G}\right) \leq E(|Y(\omega)||\mathcal{G}) \iff E\left(I_{\{\omega:|Y(\omega)|\geq a\}}|\mathcal{G}\right) \leq \frac{1}{a}E(|Y(\omega)||\mathcal{G}).$$

If we define $E\left(I_{\{\omega:|Y(\omega)|\geq a\}}|\mathcal{G}\right) := P\left(\{\omega : |Y(\omega)| \geq a\}|\mathcal{G}\right)$ we have

$$P\left(\{\omega : |Y(\omega)| \geq a\}|\mathcal{G}\right) \leq \frac{1}{a}E(|Y(\omega)||\mathcal{G}).$$

Now, to verify that $P(\cdot|\mathcal{G})$ is a legitimate probability measure note that, $E\left(I_\Omega|\mathcal{G}\right) = E\left(1|\mathcal{G}\right) = 1 = P\left(\Omega|\mathcal{G}\right)$ almost surely. Also, if $\{E_j\}_{j\in\mathbb{N}}$ is a countable collection of disjoint events $I_{\cup_{j\in\mathbb{N}}E_j} = \sum_{j\in\mathbb{N}} I_{E_j}$ and

$$P\left(\cup_{j\in\mathbb{N}}E_j|\mathcal{G}\right) = E\left(I_{\cup_{j\in\mathbb{N}}E_j}|\mathcal{G}\right) = E\left(\sum_{j\in\mathbb{N}} I_{E_j}|\mathcal{G}\right) = \sum_{j\in\mathbb{N}} E\left(I_{E_j}|\mathcal{G}\right) = \sum_{j\in\mathbb{N}} P\left(E_j|\mathcal{G}\right).$$

6. Let Y and X be random variables such that $Y, X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ and define $\varepsilon = Y - E(Y|X)$.

   (a) Show that $E(\varepsilon|X) = 0$ and $E(\varepsilon) = 0$.

   (b) Let $V(Y|X) = E(Y^2|X) - E(Y|X)^2$. Show that $V(Y|X) = V(\varepsilon|X)$, $V(\varepsilon) = E(V(Y|X))$;

   (c) $Cov(\varepsilon, h(X)) = 0$ for any function of X whose expectation exists.

   (d) Assume that $E(Y|X) = \alpha + \beta X$ where $\alpha, \beta \in \mathbb{R}$. Let $E(Y) = \mu_Y$, $E(X) = \mu_X$, $V(Y) = \sigma_Y^2$, $V(X) = \sigma_X^2$ and $\rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y}$. Show that,

$$E(Y|X) = \mu_Y + \rho\sigma_Y\frac{X - \mu_X}{\sigma_X} \text{ and } E(V(Y|X)) = (1 - \rho^2)\sigma_Y^2.$$

Answers:

(a) $E(\varepsilon|X) = E(Y - E(Y|X)|X) = E(Y|X) - E(Y|X) = 0$. By the law on iterated expectations $E(\varepsilon) = E(E(\varepsilon|X)) = 0$.

(b) $V(Y|X) = E((Y - E(Y|X))^2|X) = E(\varepsilon^2|X) = V(\varepsilon|X)$ since $E(\varepsilon|X) = 0$. Also, since $E(\varepsilon) = 0$ we have that $V(\varepsilon) = E(\varepsilon^2) = E(E(\varepsilon^2|X)) = E(V(\varepsilon|X)) = E(V(Y|X))$.

(c) $Cov(\varepsilon, h(X)) = E(\varepsilon h(X)) - E(\varepsilon)E(h(X)) = E(\varepsilon h(X))$ since $E(\varepsilon) = 0$. But by definition of conditional expectation

$$E(\varepsilon h(X)) = E(h(X)E(\varepsilon|X)) = 0 \text{ since } E(\varepsilon|X) = 0.$$

(d) First note that

$$\mu_Y = E(E(Y|X)) = E(\alpha + \beta X) = \alpha + \beta\mu_X. \tag{9.4}$$

Now, by definition of conditional expectation

$$E(XY) = E(X(\alpha + \beta X)) = \alpha\mu_X + \beta E(X^2) = \alpha\mu_X + \beta(\sigma_X^2 + \mu_X^2).$$

Also, $E(XY) = Cov(X,Y) + \mu_X\mu_Y = \rho\sigma_X\sigma_Y + \mu_X\mu_Y$. Then, we have

$$\alpha\mu_X + \beta(\sigma_X^2 + \mu_X^2) = \rho\sigma_X\sigma_Y + \mu_X\mu_Y. \tag{9.5}$$

Equations (9.4) and (9.5) form a system with two unknowns (α, β). The solution is given by,

$$\beta = \frac{\rho\sigma_Y}{\sigma_X} \text{ and } \alpha = \mu_Y - \mu_X\frac{\rho\sigma_Y}{\sigma_X}.$$

Substituting α and β into $E(Y|X) = \alpha + \beta X$ gives the desired result.

Lastly,

$$\begin{aligned}
\sigma_Y^2 := V(Y) = E\left(Y - E(Y)\right)^2 &= E\left(Y - E(Y|X) + E(Y|X) - E(Y)\right)^2 \\
&= E\left((Y - E(Y|X))^2\right) + E\left((E(Y|X) - E(Y))^2\right) \\
&+ 2E\left((Y - E(Y|X))(E(Y|X) - E(Y))\right) \\
&= E\left((Y - E(Y|X))^2\right) + V\left(E(Y|X)\right) + 2E\left(\varepsilon(E(Y|X) - E(Y))\right) \\
&= E(V(Y|X)) + V\left(E(Y|X)\right).
\end{aligned}$$

Consequently,

$$\begin{aligned}
E(V(Y|X)) &= \sigma_Y^2 - V\left(\mu_Y + \rho\sigma_Y\frac{X - \mu_X}{\sigma_X}\right) \\
&= \sigma_Y^2 - \rho^2\sigma_Y^2 = \sigma_Y^2(1 - \rho^2)
\end{aligned}$$

Chapter 10

Exercises

1. Suppose $\{X_i\}_{i=1,2,\cdots}$ is a sequence of independent and identically distributed random variables and $Y_i(x) = I_{\{\omega: X_i \le x\}}$, where I_A is the indicator function of the set A. Now define
$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i(x)$$
for fixed x. Obtain the asymptotic distribution of $\sqrt{n}(F_n(x) - F(x))$. You can use a Central Limit Theorem, but otherwise show all your work.

   Answer: (3 points) First, note that $E(Y_i(x)) = P(\{\omega : X_i \le x\}) = F(x)$ and $V(Y_i(x)) = F(x) - F(x)^2 = F(x)(1 - F(x))$.
$$\sqrt{n}(F_n(x) - F(x)) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i(x) - E(Y_i(x)))\right).$$
   Now, since the sequence is $\{Y_i(x)\}$ is IID, this is so because I_A is measurable, by Lévy's CLT
$$\frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i(x) - E(Y_i(x)))}{\sqrt{\frac{F(x)(1-F(x))}{n}}} = \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i(x) - E(Y_i(x)))}{\sqrt{F(x)(1-F(x))}} = \frac{\sqrt{n}(F_n(x) - F(x))}{\sqrt{F(x)(1-F(x))}} \xrightarrow{d} Z \sim N(0,1)$$

2. Let $\{X_n\}_{n=1,2,\cdots}$ and $\{Y_n\}_{n=1,2,\cdots}$ be sequences of random variables defined on the same probability space. Suppose $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$ and assume X_n and Y_n are independent for all n and X and Y are independent. Show that $X_n + Y_n \xrightarrow{d} X + Y$. Hint: use the characteristic function for a sum of independent random variables.

   Answer: The characteristic function of $X_n + Y_n$ is given by
$$\phi_{X_n+Y_n}(t) = E(\exp it(X_n + Y_n)) = E(\exp it(X_n)\exp it(Y_n)) = E(\exp it(X_n))E(\exp it(Y_n)) = \phi_{X_n}(t)\phi_{Y_n}($$

where the next to last equality follows by independence of X_n and Y_n. Since, $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, it must be that $\phi_{X_n}(t) \to \phi_X(t)$ and $\phi_{Y_n}(t) \to \phi_Y(t)$. So,

$$\phi_{X_n+Y_n}(t) = \phi_{X_n}(t)\phi_{Y_n}(t) \to \phi_X(t)\phi_Y(t) = \phi_{X+Y}(t),$$

where the last equality follows from independence of X and Y. Thus, $X_n+Y_n \xrightarrow{d} X+Y$.

3. Let $\{X_i\}_{i=1,2,\cdots}$ be a sequence of independent and identically random variables with $E(X_i) = 1$ and $\sigma^2_{X_i} = \sigma^2 < \infty$. Show that if $S_n = \sum_{i=1}^n X_i$

$$\frac{2}{\sigma}\left(S_n^{1/2} - n^{1/2}\right) \xrightarrow{d} Z \sim N(0,1).$$

Answer: Note that,

$$\begin{aligned} \frac{2}{\sigma}(S_n - n) &= \frac{2}{\sigma}\left(S_n^{1/2} - n^{1/2}\right)\left(S_n^{1/2} + n^{1/2}\right) \\ &= \frac{2}{\sigma}\left(S_n^{1/2} - n^{1/2}\right) n^{1/2}\left((S_n/n)^{1/2} + 1\right) \end{aligned}$$

So that,

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

and

$$\left((S_n/n)^{1/2} + 1\right)^{-1} \frac{2}{\sigma}\sqrt{n}\left((S_n/n) - 1\right) = \frac{2}{\sigma}\left(S_n^{1/2} - n^{1/2}\right).$$

Since, $\{X_i\}_{i=1,2,\cdots}$ is a sequence of independent and identically random variables with $E(X_i) = 1$, by Slutsky Theorem $\left((S_n/n)^{1/2} + 1\right)^{-1} \xrightarrow{p} 2^{-1}$ and since $\sigma^2_{X_i} = \sigma^2 < \infty$, by Lévy's CLT $\frac{1}{\sigma}\sqrt{n}\left((S_n/n) - 1\right) \xrightarrow{d} N(0,1)$. Hence, $\frac{2}{\sigma}\left(S_n^{1/2} - n^{1/2}\right) \xrightarrow{d} Z \sim N(0,1)$.