A Probability Path
Solution Manual

Sidney Resnick
School of Operations Research and Industrial Engineering
Cornell University
Rhodes Hall 214
Ithaca, NY 14853 USA

Chapter 1 Solutions

1.9.8. We have that

$$\liminf_{n\to\infty} A_n = B \cap C, \quad \limsup_{n\to\infty} A_n = B \cup C.$$

1.9.9. We write

$$A \triangle B = AB^c \bigcup BA^c$$

while

$$A^c \triangle B^c = A^c(B^c)^c \bigcup B^c(A^c)^c = A^c B \bigcup B^c A.$$

1.9.10. Suppose that $A_n \to A$. Then $\liminf_{n\to\infty} A_n = A$ so if $\omega \in A$, then for some n_0, $\omega \in A_n$, for $n \geq n_0$. Thus

$$1 = 1_A(\omega) = \lim_{n\to\infty} 1_{A_n}(\omega).$$

If $\omega \in A^c$, then $\omega \in (\limsup_{n\to\infty} A_n)^c = \liminf_{n\to\infty} A_n^c$ and therefore

$$0 = 1_A(\omega) = \lim_{n\to\infty} 1_{A_n}(\omega)$$

since $1_{A_n}(\omega) = 0$ for all sufficiently large n.

Conversely, suppose $1_{A_n} \to 1_A$. Then if $\omega \in A$, it follows that $1_{A_n}(\omega) \to 1$. Since indicator functions take on only the values 0 or 1 we get that $1_{A_n}(\omega) = 1$, for all large n, say $n \geq n_0$ and $\omega \in A_n$, for $n \geq n_0$ and $\omega \in \liminf_{n\to\infty} A_n$. Thus $A \subset \liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$.

If $\omega \in A^c$, then $1_{A_n}(\omega) \to 0$ so $\omega \in A_n^c$, for $n \geq n_0$. Hence $A^c \subset \liminf_{n\to\infty} A_n^c$, or equivalently $A \supset (\liminf_{n\to\infty} A_n^c)^c = \limsup_{n\to\infty} A_n$.

1.9.11. We first show that

$$\bigcup_n [0, a_n) = [0, \sup_n a_n).$$

If ω belongs to the left side union, then $\omega < a_n$ for some n and therefore $\omega < \sup_n a_n$ and $\omega \in [0, \sup_n a_n)$ which is the right side. If $\omega \in [0, \sup_n a_n)$, that is, ω belongs to the interval on the right, then $\omega < \sup_n a_n$ and $\omega < a_n$ for some n which implies $\omega \in \cup_n [0, a_n)$.

For the second part, $\sup_n \frac{n}{n+1} = 1$ and $1 \notin \cup_n [0, \frac{n}{n+1}] = [0, 1)$.

1.19.14. Suppose $\mathcal{A}_n$ is a field for each n and that $\mathcal{A}_n \uparrow$. Since $\mathcal{A}_n$ is a field $\Omega \in \mathcal{A}_n$ for all n and therefore $\Omega \in \cup_n \mathcal{A}_n$. If $A \in \cup_n \mathcal{A}_n$, then $A \in \mathcal{A}_n$ for some n which implies $A^c \in \mathcal{A}_n$ which implies $A^c \in \cup_n \mathcal{A}_n$. So $\cup_n \mathcal{A}_n$ is closed under complementation.

If $A, B \in \cup_n \mathcal{A}_n$, there exist n, m such that $A \in \mathcal{A}_n$ and $B \in \mathcal{A}_m$. Thus $A, B \in \mathcal{A}_{n \vee m}$ and $A \cap B \in \mathcal{A}_{n \vee m}$ (since fields are closed under finite intersection). This yields $AB \in \cup_n \mathcal{A}_n$.

1.9.15. We suppose $\Omega = \{1, 2, ...\}$ and define

$$\mathcal{C}_j = \{\Lambda : \Lambda \subset \{1, 2, ...j\}\}.$$

Set $\sigma(\mathcal{C}_j) =: \mathcal{B}_j$. Check that

$$\mathcal{B}_j = \mathcal{C}_j \cup \{\Lambda \cup \{j+1, j+2, ...\} : \Lambda \in \mathcal{C}_j\}.$$

If $\Lambda \in \mathcal{B}_j$ satisfies the property that the number of elements of Λ is infinite, then $\Lambda \supset \{j+1, j+2, ...\}$.

Let $A_j = \{2j-1\}$, $j = 1, 2, ...$ so

$$(A_1, A_2, A_3, \ldots) = (\{1\}, \{3\}, \{5\}, \ldots).$$

Then

$$A_j \in \mathcal{B}_{2j-1} \subset \bigcup_n \mathcal{B}_n$$

but

$$\bigcup_j A_j = \{1, 3, 5, 7, \ldots\} \notin \bigcup_n \mathcal{B}_n,$$

since $\cup_j A_j$ is an infinite set but for no j is it true that $\{1, 3, 5, 7, \ldots\} \supset \{j+1, j+2, ...\}$.

Note that a union of σ-fields is not necessarily even a field. Let $\Omega = \{1, 2, \ldots\}$ and $\mathcal{B}_i = \sigma(\{i\}) = \{\emptyset, \Omega, \{i\}, \{i\}^c\}$ for $i = 1, 2$. Then $\{i\} \in \mathcal{B}_i$ but $\{1\} \cup \{2\} = \{1, 2\} \notin \mathcal{B}_1 \cup \mathcal{B}_2$.

1.9.17. We have $\omega \in \liminf_{n\to\infty} A_n$ iff $\omega \in A_n$ for $n \geq n_0$ for some n_0. This is equivalent to $1_{A_n}(\omega) = 1$ for $n \geq n_0$. But since indicators only take values 0 or 1, the only way a sequence of indicators can converge to a limit is if the indicators equal the limit from some index on. This means that the statement: for some $n \geq n_0$, $1_{A_n}(\omega) = 1$ is equivalent to $\lim_{n\to\infty} 1_{A_n}(\omega) = 1$.

1.9.18. We check the three postulates for a field or algebra:

(i) $\Omega \in \mathcal{A}$ by assumption

(ii) Complementation: If $A \in \mathcal{A}$, then since $\Omega \in \mathcal{A}$ we have $\Omega A^c = A^c \in \mathcal{A}$.

(iii) Suppose $A, B \in \mathcal{A}$. Then $AB^c \in \mathcal{A}$ so $A \cap (AB^c)^c = A \cap (A^c \cup B) = AA^c \cup AB = AB \in \mathcal{A}$.

1.9.19. We have $1_{A\cup B}(\omega) = 1$ iff $\omega \in A \cup B$ iff either $1_A(\omega) = 1$ or $1_B(\omega) = 1$ iff $1_A(\omega) \vee 1_B(\omega) = 1$.

Likewise, $1_{A\cap B}(\omega) = 1$ iff $\omega \in A \cap B$ iff both $1_A(\omega) = 1$ and $1_B(\omega) = 1$ iff $1_A(\omega) \wedge 1_B(\omega) = 1$.

Since indicators take only values 0 or 1 we are done.

1.9.20. Define

$$\Lambda := \{\sum_{i=1}^{m} \bigcap_{j=1}^{n_i} A_{ij} : A_{ij} \in \mathcal{C} \text{ or } A_{ij}^c \in \mathcal{C}\}$$

and remember the summation notation for sets implies a *disjoint* union. We claim Λ is a field and verify the field postulates:

(i) Pick $A \in \mathcal{C}$ so that $(A^c)^c \in \mathcal{C}$ and thus

$$\Omega = A + A^c \in \Lambda.$$

(iii) Closure under finite intersection: Suppose

$$\sum_{i\in I} \bigcap_{j\in J_i} A_{ij} \text{ and } \sum_{k\in I'} \bigcap_{l\in J'_k} A'_{kl}$$

are two sets in Λ. Then the intersection is

$$\sum_{(i,k)\in I\times I'} \left(\bigcap_{j\in J_i} A_{ij} \bigcap \bigcap_{l\in J'_k} A'_{kl} \right)$$

which is also in Λ.

(ii) Closure under complementation: The complement of a typical set in Λ is

$$\left(\sum_{i=1}^{m} \bigcap_{j=1}^{n_i} A_{ij} \right)^c = \bigcap_{i=1}^{m} \bigcup_{j=1}^{n_i} A_{ij}^c.$$

To show that this set is in Λ, it suffices because of (iii) just checked, to verify that one of the sets in the intersection is in Λ and hence it suffices to show that $\cup_{j=1}^{n} A_j^c \in \Lambda$ where $A_j \in \mathcal{C}$ or $A_j^c \in \mathcal{C}$. However, we may write

$$\bigcup_{j=1}^{n} A_j^c = A_1^c + A_2^c A_1 + A_3^c A_1 A_2 + \cdots + A_n^c A_1 A_2 \ldots A_{n-1},$$

which is a disjoint sum of sets of the form $\cap_{i=1}^{k} B_i$ where $B_i \in \mathcal{C}$ or $B_i^c \in \mathcal{C}$. Therefore $\bigcup_{j=1}^{n} A_j^c \in \Lambda$ as required.

So Λ is a field. For any $A \in \mathcal{C}$, $A \in \Lambda$ so $\mathcal{C} \subset \Lambda$ and therefore the minimal field over $\mathcal{C}$ is contained in Λ:

$$\mathcal{A}(\mathcal{C}) \subset \Lambda.$$

Also, if A_{ij} or $A_{ij}^c \in \mathcal{C}$, then $A_{ij} \in \mathcal{A}(\mathcal{C})$. Therefore $\cap_{j=1}^{n_i} A_{ij} \in \mathcal{A}(\mathcal{C})$ so $\sum_{i=1}^{m} \cap_{j=1}^{n_i} A_{ij} \in \mathcal{A}(\mathcal{C})$. We conclude that $\Lambda \subset \mathcal{A}(\mathcal{C})$ as well.

1.9.26a. Set $\mathcal{C} = \{A_1, ..., A_n\}$, where $\sum_{i=1}^{k} A_i = \Omega$. We claim

$$\mathcal{A}(\mathcal{C}) = \{\bigcup_{i \in I} A_i, \quad I \subset \{1, ..., k\}\}$$

is the minimal algebra over $\mathcal{C}$. Denote the right side collection of sets by $\mathcal{A}$. To prove the claim, we first show that $\mathcal{A}$ is a field. To do this, we verify the postulates.

(i) First of all, $\Omega \in \mathcal{A}$ since we may take $I = \{1, ..., n\}$.

(ii) If $A = \cup_{i \in I} A_i \in \mathcal{A}$, then $A^c = \cup_{i \in I^c} A_i \in \mathcal{A}$.

(iii) If $A_j = \cup_{i \in I_j} A_i$, for $j = 1, 2$, then $A_1 \cup A_2 = \cup_{i \in I_1 \cup I_2} A_i \in \mathcal{A}$.

So $\mathcal{A}$ is a field, $\mathcal{A} \supset \mathcal{C}$, so by minimality we have $\mathcal{A} \supset \mathcal{A}(\mathcal{C})$. But clearly, since $A_i \in \mathcal{C} \subset \mathcal{A}(\mathcal{C})$, we have $\mathcal{A} \subset \mathcal{A}(\mathcal{C})$. The two set inclusions give the desired equality.

1.9.27. Call $\mathbb{Q}$ the rational numbers and define

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a, b] : -\infty \le a \le b < \infty\}$$

and

$$\mathcal{F} = \sigma\{(a, b] : -\infty \le a \le b < \infty,\ a, b \in \mathbb{Q}\}.$$

For $q, s \in \mathbb{Q}$,

$$(q, s] \in \{(a, b] : -\infty \le a \le b < \infty\} \subset \mathcal{B}(\mathbb{R}).$$

Therefore $\mathcal{F} \subset \mathcal{B}(\mathbb{R})$.

On the other hand, for any a, b

$$(a, b] = \lim_{n \to \infty} (q_n, s_n]$$

where $q_n \downarrow a$ and $s_n \downarrow b$ and $q_n, s_n \in \mathbb{Q}$. So $(a, b] \in \mathcal{F}$ and $\mathcal{B}(\mathbb{R}) \subset \mathcal{F}$.

1.9.28. Let $\mathbb{Z} = \{..., -1, 0, 1, 2, ...\}$. Let $\mathcal{F}$ be the periodic sets. A set A is *periodic*, written $A \in \mathcal{F}$, if for all natural numbers $n \in \mathbb{Z}$ we have $x \in A$ iff $x \pm n \in A$. We verify the σ-field postulates for $\mathcal{F}$:

(a) First of all, $\mathbb{R} \in \mathcal{F}$.

(b) Complementation: Next, suppose $A \in \mathcal{F}$ and we show $A^c \in \mathcal{F}$. If $x \in A^c$, then for any $n \in \mathbb{Z}$, we claim $x+n \in A^c$. If not then $x+n \in A$ and since A is periodic $(x+n) - n = x \in A$, a contradiction.

(c) Closure under countable unions: Let $B_j \in \mathcal{F}$ for $j \geq 1$. We show $\cup_j B_j \in \mathcal{F}$. If $x \in \cup_j B_j$, then there exists j_0 such that $x \in B_{j_0}$. For any $n \in \mathbb{Z}$, $x + n \in B_{j_0} \subset \cup_j B_j$.

1.9.29. Let $\mathcal{D}(\mathcal{C})$ be the smallest class containing $\mathcal{C}$ and closed under countable intersection and union. This minimal structure exists since closure axioms define the structure. Then $\mathcal{C} \subset \mathcal{D}(\mathcal{C})$. Also $\sigma(\mathcal{C})$ is closed under countable union and intersection and since $\sigma(\mathcal{C}) \supset \mathcal{C}$, we get

$$\sigma(\mathcal{C}) \supset \mathcal{D}(\mathcal{C}).$$

Let

$$\mathcal{F} := \{\Lambda \in \mathcal{D}(\mathcal{C}) : \Lambda^c \in \mathcal{D}(\mathcal{C})\}.$$

We claim $\mathcal{F}$ is a σ-field. Note if $\Lambda_n \in \mathcal{F}$, then $\Lambda_n \in \mathcal{D}(\mathcal{C})$ and $\Lambda_n^c \in \mathcal{D}(\mathcal{C})$. This means $\cup_n \Lambda_n \in \mathcal{D}(\mathcal{C})$ and therefore

$$\Big(\bigcup_n \Lambda_n\Big)^c = \bigcap_n \Lambda_n^c \in \mathcal{D}(\mathcal{C}),$$

since $\Lambda_n^c \in \mathcal{D}(\mathcal{C})$. So $\mathcal{F}$ is closed under countable unions.

If $\Lambda \in \mathcal{F}$ so $\Lambda^c \in \mathcal{D}(\mathcal{C})$, then Λ^c satisfies

$$(\Lambda^c)^c = \Lambda \in \mathcal{F} \subset \mathcal{D}(\mathcal{C})$$

which implies $\Lambda^c \in \mathcal{F}$. So $\mathcal{F}$ is closed under complements.

Is $\Omega \in \mathcal{F}$? Since $\Lambda \in \mathcal{D}(\mathcal{C})$ implies $\Lambda^c \in \mathcal{D}(\mathcal{C})$ and $\Omega = \Lambda + \Lambda^c \in \mathcal{D}(\mathcal{C})$ and $\emptyset = \Lambda \cap \Lambda^c \in \mathcal{D}(\mathcal{C})$, we get $\Omega \in \mathcal{F}$.

We claim, next, that $\mathcal{F} \supset \mathcal{C}$. The reason for this is that if $\Lambda \in \mathcal{C} \subset \mathcal{D}(\mathcal{C})$, then

$$\Lambda^c = \bigcup_i C_i \in \mathcal{D}(\mathcal{C})$$

where $\{C_i\}$ are each sets in $\mathcal{C}$. So $\mathcal{F}$ is a σ-field, $\mathcal{F} \supset \mathcal{C}$, so $\mathcal{F} \supset \sigma(\mathcal{C})$. But by definition, $\mathcal{F} \subset \mathcal{D}(\mathcal{C})$. We conclude that $\sigma(\mathcal{C}) \subset \mathcal{D}(\mathcal{C})$.

1.9.31. If Ω is countable, then $\mathcal{C} := \{\{x\} : x \in \Omega\}$ is a countable generating class, since for any $A \subset \Omega$, $A = \cup_{\alpha \in A} A_\alpha$.

Now let Ω be uncountable. For the purpose of getting a contradiction, suppose $\mathcal{C} = \{C_n, n \geq 1\}$ is a countable generating class for the σ-field of countable-cocountable sets. Define

$$C_n^{\#} = \begin{cases} C_n, & \text{if } C_n \text{ is countable,} \\ C_n^c, & \text{otherwise.} \end{cases}$$

So $C_n^{\#}$ is always countable and so is $C = \cup_n C_n^{\#}$. Therefore, C^c is uncountable.

Pick $x, y \in C^c$ such that $x \neq y$. For any n,

$$\{x, y\} \subset \begin{cases} C_n, & \text{if } C_n \text{ is not countable,} \\ C_n^c, & \text{otherwise.} \end{cases}$$

Let

$$\mathcal{F} = \{A \in \sigma(\mathcal{C}) : \{x, y\} \subset A \text{ or } \{x, y\} \subset A^c\}.$$

For any n, $\{x, y\} \subset$ either C_n or C_n^c so $\mathcal{C} \subset \mathcal{F}$. Further properties of $\mathcal{F}$:

1. $\{x, y\} \subset \Omega$ so $\Omega \in \mathcal{F}$.
2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
3. If $A_n \in \mathcal{F}$, then since $\mathcal{F} \subset \sigma(\mathcal{C})$, $\cup_n A_n \in \sigma(\mathcal{C})$. If there exists n such that $\{x, y\} \subset A_n$, then $\{x, y\} \subset \cup_n A_n$ and $\cup_n A_n \in \mathcal{F}$. If for all n, $\{x, y\} \subset A_n^c$, then $\{x, y\} \subset \cap_n A_n^c$ which implies $\cup_n A_n \in \mathcal{F}$.

So we conclude $\mathcal{F}$ is a σ-field and since $\mathcal{C} \subset \mathcal{F}$, we get $\sigma(\mathcal{C}) \subset \mathcal{F}$ and since also $\mathcal{F} \subset \sigma(\mathcal{C})$ we get $\mathcal{F} = \sigma(\mathcal{C})$.

For $A \in \mathcal{F}$, either $\{x, y\} \subset A$ or $\{x, y\} \subset A^c$. But $\{x\} \in \sigma(\mathcal{C})$ and $\{x, y\} \not\subset \{x\}$ and $\{x, y\} \not\subset \{x\}^c$. So $\{x\} \notin \mathcal{F}$ which gives a contradiction.

1.9.34. Let

$$\mathcal{G} := \{AB + A^c B' : \; B, B' \in \mathcal{B}\}.$$

We claim $\mathcal{G}$ is a σ-field. To verify this note

1. $\Omega = A\Omega + A^c\Omega \in \mathcal{G}$.
2. If $B_n, B_n' \in \mathcal{B}$ for $n \geq 1$, then $AB_n + A^c B_n'$ and

$$\bigcup_n AB_n + A^c B_n' = A \cap \Big(\bigcup_n B_n\Big) + A^c \cap \Big(\bigcup_n B_n'\Big) \in \mathcal{G}$$

   since $\cup_n B_n$ and $\cup_n B_n'$ are both in $\mathcal{B}$.

3. If $AB + A^c B' \in \mathcal{G}$ then

$$\begin{aligned}
(AB + A^c B')^c =& (AB)^c \cap (A^c B')^c = (A^c \cup B^c) \cap (A \cup (B')^c) \\
=& \Big(A^c (B')^c + AB^c\Big) \cup B^c (B')^c \\
=& A^c (B')^c \cup AB^c \cup AB^c (B')^c \cup A^c B^c (B')^c \\
=& A\big(B^c \cup B^c (B')^c\big) \cup A^c\big((B')^c \cup B^c (B')^c\big) \\
=& AB^c + A^c (B')^c.
\end{aligned}$$

So $\mathcal{G}$ is a σ-algebra.

Also, we have $A \in \mathcal{G}$ and $\mathcal{B} \subset \mathcal{G}$ and therefore

$$\mathcal{G} \supset \sigma(\mathcal{B}, A).$$

Also,

$$\mathcal{G} \subset \sigma(\mathcal{B} \cup \{A\})$$

since the right side contains $\mathcal{B}$ and A and hence contains sets of the form $BA + B'A^c$.

1.9.35. Suppose $\mathcal{F}$ is a countably-infinite σ-field so that we can write it as

$$\mathcal{F} = \{B_1, B_2, \dots\}.$$

For $\mathbb{N} = \{1, 2, \dots\}$, let

$$\epsilon = (\epsilon_1, \epsilon_2, \dots) \in \{0, 1\}^{\mathbb{N}},$$

and write

$$B_\epsilon = \bigcap_{i=1}^{\infty} B_i^{\epsilon_i},$$

where

$$B_i^{\epsilon_i} = \begin{cases} B_i, & \text{if } \epsilon_i = 1, \\ B_i^c, & \text{if } \epsilon_i = 0. \end{cases}$$

Set

$$\mathcal{C} = \left\{B_\epsilon, \epsilon \in \{0, 1\}^{\mathbb{N}}\right\}.$$

Since $\mathcal{F}$ is a σ-field, $\mathcal{C} \in \mathcal{F}$. Note also that

$$B_\epsilon \bigcap B_{\epsilon'} = \emptyset, \quad \text{if } \epsilon \neq \epsilon',$$

so sets of $\mathcal{C}$ partition Ω.

Now we *claim* that $\mathcal{C}$ contains infinitely many non-empty sets. If not, then there are finitely many non-empty sets in $\mathcal{C}$ which partition Ω. This implies $\sigma(\mathcal{C})$ is finite. But $\sigma(\mathcal{C}) = \mathcal{F}$ since

(a) $\mathcal{C} \subset \mathcal{F}$ implies $\sigma(\mathcal{C}) \subset \mathcal{F}$.

(b) If $B_k \in \mathcal{F}$, then

$$B_k = \bigcup_{\epsilon : \epsilon_k = 1} B_\epsilon \in \sigma(\mathcal{C}),$$

and hence $\mathcal{F} \subset \sigma(\mathcal{C})$.

This would mean that $\mathcal{F}$ is finite which contradicts the assumption that $\mathcal{F}$ is countably infinite.

So, since $\mathcal{C}$ has infinitely many non-empty sets, we write

$$\mathcal{C} := \{\emptyset, C_1, C_2, \dots\}$$

where $C_i \neq \emptyset$, $i \geq 1$.

Define a function f on the subsets of $\mathbb{N}$ by

$$f : \mathcal{P}(\mathbb{N}) \mapsto \mathcal{F}, \quad f(I) = \bigcup_{i \in I} C_i.$$

We claim that f is 1-1. To see this note that $\cup_{i\in I} C_i = \cup_{j \in I'} C_j$ implies that $I = I'$ since The C_i's are disjoint and non-empty.

So for all $I \in \mathcal{P}(\mathbb{N})$, $f(I) \in \mathcal{F}$ and hence $\mathcal{F}$ cannot be countably infinite since a subset $\{f(I), I \in \mathcal{P}(\mathbb{N})$ is in 1-1 correspondence with $\mathcal{P}(\mathbb{N})$ which has cardinality $2^{\aleph_0}$.

1.9.44. To see that $\mathcal{A} \subset \bar{\mathcal{A}}$, note that if $A \in \mathcal{A}$, then we may set $A_n = A$ so that $A_n \to A$ showing that $A \in \bar{\mathcal{A}}$.

We now see why $\bar{\mathcal{A}}$ is a field. We verify the field postulates:

1. Since $\emptyset, \Omega \in \mathcal{A}$, and $\mathcal{A} \subset \bar{\mathcal{A}}$ we have $\emptyset, \Omega \in \bar{\mathcal{A}}$.

2. Suppose $A \in \bar{\mathcal{A}}$. Then there exist $A_n \in \mathcal{A}$ and $A_n \to A$. Since $\mathcal{A}$ is a field, we have $A_n^c \in \mathcal{A}$. Thus

$$\limsup_{n\to\infty} A_n^c = \left(\liminf_{n\to\infty} A_n\right)^c = A^c,$$
$$\liminf_{n\to\infty} A_n^c = \left(\limsup_{n\to\infty} A_n\right)^c$$

   and so $A_n^c \to A^c$. Thus $A \in \bar{\mathcal{A}}$.

3. Suppose $A, B \in \bar{\mathcal{A}}$. Then there exist $A_n \in \mathcal{A}$, $B_n \in \mathcal{A}$ such that

$$A_n \to A, \quad B_n \to B.$$

   It follows that $A_n B_n \in \mathcal{A}$ and we show that $A_n B_n \to AB$ proving that $\bar{\mathcal{A}}$ is closed under finite intersections. First of all

$$\limsup_{n\to\infty} A_n \bigcap B_n = \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} A_k B_k \subset \limsup_{n\to\infty} A_n = A$$

   and similarly

$$\limsup_{n\to\infty} A_n \bigcap B_n = \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} A_k B_k \subset \limsup_{n\to\infty} B_n = B$$

so that

$$\limsup_{n\to\infty} A_n \bigcap B_n \subset AB.$$

On the other hand, since $\liminf_{n\to\infty} A_n B_n$ is the points in $A_n B_n$ for all large n, we have

$$\liminf_{n\to\infty} A_n B_n = \liminf_{n\to\infty} A_n \bigcap \liminf_{n\to\infty} B_n = AB.$$

Thus

$$AB = \liminf_{n\to\infty} A_n B_n \subset \limsup_{n\to\infty} A_n B_n \subset AB.$$

Thus $A_n B_n \to AB$ and $AB \in \bar{\mathcal{A}}$.

Chapter 2 Solutions

2.6.1.

(a) First of all $\Omega^c = \emptyset$ is finite so $\Omega \in \mathcal{F}_0$.

Next check closure under complementation: If $A \in \mathcal{F}_0$ then either A or A^c is finite. Therefore $A^c \in \mathcal{F}_0$ since either $(A^c)^c$ or A^c is finite.

Finally check closure under finite intersection: Suppose $A_i \in \mathcal{F}_0$, $i = 1, 2$. If one of A_1, A_2 is finite, then A_1A_2 is finite and hence in $\mathcal{F}_0$. If neither set is finite, then A_1^c and A_2^c are finite, so $A_1^c \cup A_2^c$ is finite. Therefore $(A_1^c \cup A_2^c)^c = A_1A_2 \in \mathcal{F}_0$.

(b) Let $E_1, ..., E_k \in \mathcal{F}_0$, $E_i \cap E_j = \emptyset$, for $i \neq j$. At most one can be infinite, since if E_1 and E_2 are both infinite and $E_1 \cap E_2 = \emptyset$, then E_1^c, E_2^c are finite which implies $E_1^c \cup E_2^c$ is finite. So $(E_1^c \cup E_2^c)^c$ is infinite and in $\mathcal{F}_0$. However, we also have $(E_1^c \cup E_2^c)^c = E_1E_2 = \emptyset$, which gives a contradiction.

If none of $E_1, E_2, \ldots E_k$ is infinite then

$$P(\bigcup_{j=1}^{k} E_j) = 0 = \sum_{j=1}^{k} P(E_j),$$

If exactly one is infinite, then $\bigcup_{j=1}^k E_j$ is infinite and $P(\bigcup_{j=1}^k E_j) = 1 = \sum_{j=1}^{k} PE_j$, since the latter is a sum of $(k-1)$ zeros and one 1.

P is not σ-additive. Let Ω_N be finite and $\Omega_N \uparrow \Omega$. If P were σ-finite, we would have

$$0 = P(\Omega_N) \uparrow P(\Omega) = 1.$$

(c) Define

$$P(E) = \begin{cases} 0, & \text{if } E \text{ is finite,} \\ 1, & \text{if } E^c \text{ is finite.} \end{cases}$$

Suppose $E_1, E_2, \cdots \in \mathcal{F}_0$ and $\bigcup_i E_i \in \mathcal{F}_0$ and $\{E_n\}$ are mutually disjoint. As in (b), at most one E_n can be infinite. Then either

(I) $\bigcup_i E_i$ is finite, in which case E_i is finite for all i and $P(\bigcup_i E_i) = 0 = \sum_i P(E_i)$

or

(II) $(\bigcup_i E_i)^c$ is finite. This means there exists i such that E_i^c is finite and and because at most one of the $\{E_n\}$ can be infinite, for all $j \neq i$, E_j is finite. Therefore

$$P(\bigcup_i E_i) = 1 = \sum_k P(E_k) = P(E_i) + \sum_{j \neq i} P(E_i) = 1 + 0.$$

2.6.2. The result can be proven using the representation for $\mathcal{A}$ (see Problem 1.9.20, page 23)

$$\mathcal{A} = \{\bigcup_{i=1}^{m}\bigcap_{j=1}^{n_i} A_{ij} : A_{ij} \in \mathcal{P} \text{ or } A_{ij}^c \in \mathcal{P} \text{ and } \bigcap_{j=1}^{n_i} A_{ij},\ i = 1, \ldots, m \text{ are disjoint}\}.$$

Given two probability measures P_1 and P_2 which agree on $\mathcal{P}$ we need

$$P_1(\bigcup_{i=1}^{m}\bigcap_{j=1}^{n_i} A_{ij}) = \sum_{i=1}^{m} P_1\Big(\bigcap_{j=1}^{n_i} A_{ij}\Big)$$

to be equal to

$$= \sum_{i=1}^{m} P_2\Big(\bigcap_{j=1}^{n_i} A_{ij}\Big).$$

Therefore, it suffices to prove for $A_1, \ldots, A_k$, where $A_l \in \mathcal{P}$ or $A_l^c \in \mathcal{P}$, $l = 1, \ldots, k$ that

$$P_1(\bigcap_{l=1}^{k} A_l) = P_2(\bigcap_{l=1}^{k} A_l).$$

Separate the A's into two groups $\{A_i, i \in I\}$ and $\{A_l, l \in J\}$ where $I + J = \{1, \ldots, k\}$ and $A_i \in \mathcal{P}$ for $i \in I$ and $A_l^c \in \mathcal{P}$ for $l \in J$. Call $B_1 = \cap_{i \in I} A_i$ so that $B_1 \in \mathcal{P}$ since $\mathcal{P}$ is a π-system. We need to prove

$$P_1(B_1 \bigcap (\bigcap_{l \in J} A_l)) = P_2(B_1 \bigcap (\bigcap_{l \in J} A_l)).$$

Write

$$\begin{aligned} P_i(B_1 \bigcap \bigcap_{l \in J} (A_l)) &= P_i(B_1 \bigcap (\bigcup_{l \in J} A_l^c)^c) = P_i(B_1) - P_i(B_1 \bigcup_{l \in J} A_l^c) \\ &= P(B_1) - P(\bigcup_{l \in J} B_1 A_j^c) \end{aligned}$$

and apply inclusion-exclusion.

2.6.3. If $B_i \subset A_i$ then $\cup_i B_i \subset \cup_i A_i$ and

$$\bigcup_i A_i \setminus \bigcup B_i = \Big(\bigcup_i A_i\Big) \bigcap \Big(\bigcup_i B_i\Big)^c = \Big(\bigcup_i A_i\Big) \bigcap \Big(\bigcap_i B_i^c\Big)$$
$$= \bigcup_i \Big(A_i \bigcap_j B_j^c\Big) \subset \bigcup_i A_i B_i^c$$

so

$$P\Big(\bigcup_i A_i\Big) - P\Big(\bigcup_i B_i\Big) \le P\Big(\bigcup_i A_i B_i^c\Big) \le \sum_i P(A_i B_i^c) = \sum_i (P(A_i) - P(B_i)).$$

2.6.4. First of all, the extension is certainly not unique. For an easy example, take $\mathcal{B} = \{\emptyset, \Omega\}$ and $A \notin \mathcal{B}$. Then

$$\mathcal{B}_1 = \sigma(A, \mathcal{B}) = \{\emptyset, \Omega, A, A^c\}.$$

Knowing a probability on $\{\emptyset, \Omega\}$ does not give much instruction about how to extend it to A and A^c.

Here is one way to extend using outer and inner measure. For any $S \subset \Omega$, define

$$P^*(S) := \inf\{P(B) : S \subset B, B \in \mathcal{B}\}.$$

Let $B_n \in \mathcal{B}$, $S \subset B_n$, $P(B_n) \downarrow P^*(S)$. Such a sequence $\{B_n\}$ exists by definition of "inf". Now define

$$S^* = \bigcap_{n=1}^{\infty} B_n = \lim_{N\to\infty} \bigcap_{n=1}^{N} B_n.$$

Thus, $S^* \in \mathcal{B}$, $S \subset S^*$, and therefore

$$P^*(S) \le P(S^*)$$

(from the definition of P^*)

$$= \lim_{N\to\infty} \downarrow P\Big(\bigcap_{n=1}^{N} B_n\Big) \le \lim_{N\to\infty} P(B_N) = P^*(S).$$

We conclude

$$P^*(S) = P(S^*). \tag{2.6.4.1}$$

Next, we claim, if $C \in \mathcal{B}$ and

$$\mathcal{B} \ni C \subset S^* \setminus S, \text{ then } P(C) = 0. \tag{2.6.4.2}$$

This follows from $S \subset S^* \setminus C$, $S^* \setminus C \in \mathcal{B}$, and thus

$$P(S^*) = P^*(S) \le P(S^* \setminus C)$$

(from the definition of P^*)

$$= P(S^*) - P(C),$$

(since $C \subset S^*$) whence $P(C) = 0$.

Next define, for any $S \subset \Omega$,

$$S_* = \left((S^c)^*\right)^c,$$

so that $S_* \in \mathcal{B}$, $S_*^c = (S^c)^* \supset S^c$ which yields, by taking inverses $S_* \subset S$. Then

$$\begin{aligned} P(S_*) =& 1 - P\left((S^c)^*\right) = 1 - P^*(S^c) \\ =& 1 - \inf\{P(\Lambda) : S^c \subset \Lambda, \Lambda \in \mathcal{B}\} \\ =& \sup\{P(\Lambda^c) : S^c \subset \Lambda, \Lambda \in \mathcal{B}\} \\ =& \sup\{P(V) : S^c \subset V^c, V \in \mathcal{B}\} \\ =& \sup\{P(V) : S \supset V, V \in \mathcal{B}\}. \end{aligned}$$

Also, as with (2.6.4.2), if $D \subset S \setminus S_*$, and $D \in \mathcal{B}$, then $P(D) = 0$.

Pick $\lambda \in [0,1]$ and define P_1 on $\mathcal{B}_1 = \sigma(\mathcal{B}, A) = \{BA \bigcup B'A^c; B, B' \in \mathcal{B}\}$ by

$$\begin{aligned} P_1(BA \bigcup B'A^c) :=& \lambda P(A^*B) + (1-\lambda)P(A_*B) \\ & \lambda P\left((A^*)^c B'\right) + (1-\lambda)P\left((A_*)^c B'\right), \end{aligned}$$

so that

$$P_1(BA \bigcup B'A^c) = P_1(BA) + P_1(B'A^c).$$

Here are the relevant properties of P_1:

1. P_1 is well defined on $\mathcal{B}_1 = \sigma(\mathcal{B}, A)$.

2. P_1 extends P. This is clear since if $B \in \mathcal{B}$,

$$\begin{aligned} P_1(BA \bigcup B'A^c) =& P_1(B) \\ =& \lambda \left[P(A^*B) + P\left((A^*)^c B\right)\right] \\ & (1-\lambda)\left[P(A_*B) + P\left((A_*^c)B\right)\right] \\ =& \lambda P(B) + (1-\lambda)P(B) = P(B). \end{aligned}$$

3. P_1 is a probability measure.

To see why P_1 is a probability measure, note that clearly $P_1(C) \geq 0$ for all $C \in \mathcal{B}_1$ and $P_1(\Omega) = P(\Omega) = 1$. To verify σ-additivity, suppose $B_n A \cup B_n' A^c \in \mathcal{B}_1$ are disjoint for $n \geq 1$, where $B_n, B_n' \in \mathcal{B}_1$. This means that $\{AB_n, n \geq 1\}$ are disjoint and $\{A^c B_n', n \geq 1\}$ are disjoint.

For any $n \neq m$, $AB_n \cap AB_m = \emptyset$ implies $\emptyset = A_*(AB_n \cap AB_m) = A_*B_n \cap A_*B_m$ since $A_* \subset A$. Also $A^*B_n \cap A^*B_m \subset A^*$ but

$$(A^*B_n \bigcap A^*B_m) \cap A = A^* \bigcap (B_nA \cap B_mA) = \emptyset,$$

so

$$A^*B_n \bigcap A^*B_m \subset A^* \setminus A$$

and by (2.6.4.2)

$$P(A^*B_n \bigcap A^*B_m) = 0.$$

So $\{A^*B_n, n \geq 1\}$ are almost disjoint (see Problem 2.6.6) and

$$\begin{aligned}
P_1(\bigcup_{n=1}^{\infty} AB_n) =& \lambda P_1\Big((\bigcup_{n=1}^{\infty} B_n)A\Big) \\
=& \lambda P\Big((\bigcup_n B_n)A^*\Big) + (1-\lambda)P\Big((\bigcup_n B_n)A_*\Big) \\
=& \lambda P\Big(\bigcup_n (B_nA^*)\Big) + (1-\lambda)P\Big(\bigcup_n (B_nA_*)\Big) \\
=& \lambda \sum_n P(B_nA^*) + (1-\lambda)\sum_n P(B_nA_*) \\
=& \sum_n [\lambda P(B_nA^*) + (1-\lambda)P(B_nA_*)] \\
=& \sum_n P_1(B_nA).
\end{aligned}$$

A similar argument works on A^c.

We conclude that P_1 is a probability measure that extends P on $\mathcal{B}$ to $\mathcal{B}_1$.

2.6.6. Since

$$P(\cup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(\cup_{j=1}^{n} A_j)$$

and

$$\sum_{j=1}^{\infty} PA_j = \lim_{n\to\infty} \sum_{j=1}^{n} PA_j,$$

it suffices to show $P(\bigcup_{j=1}^{n} A_j) = \sum_{j=1}^{n} PA_j$. To check this observe that by the Bonferroni inequality

$$\sum_{j=1}^{n} PA_j - \sum_{1\leq i<j\leq n} P(A_iA_j) \leq P(\bigcup_{j=1}^{n} A_j) \leq \sum_{1}^{n} P(A_j)$$

and since $\sum_{i<j} P(A_iA_j) = 0$, the result follows.

2.6.8. We summarize the probabilities in the following chart

Ω	a	b	c	d
P_1	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
P_2	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$.

Set

$$\mathcal{C} = \{\{a, b\}, \{d, c\}, \{a, c\}, \{b, d\}\}$$

and note $\mathcal{C}$ is not a π-system since

$$\{a, b\} \bigcap \{a, c\} = \{a\} \notin \mathcal{C}.$$

Check that $\sigma(\mathcal{C}) = \mathcal{P}(\Omega)$ and $P_1 = P_2$ on $\mathcal{C}$ but not on $\sigma(\mathcal{C})$ since, for instance, $P_1(\{a\}) = \frac{1}{6} \neq \frac{1}{3} = P_2(\{a\})$.

2.6.9. (a) First of all, if $F(x) - F(x-) > 0$, then $P\{x\} > 0$ and if $B \subset \{x\}$, then either $B = \emptyset$ so that $P(B) = P(\emptyset) = 0$ or $B = \{x\}$ so $P(B) = P\{x\}$. Thus

$$\{x : F(x) - F(x-) > 0\} \subset \{\text{ atoms of } P\}.$$

Next suppose A is an atom of P. Define

$$\alpha := \sup\{x : P((-\infty, x) \cap A) = 0\}, \quad \beta := \inf\{x : P((x, \infty) \cap A) = 0\}.$$

If $\alpha < \beta$, then $P((\alpha, \infty + \frac{\beta-\alpha}{3}) \cap A) > 0$, and $P((\beta - \frac{\beta-\alpha}{3}, \beta) \cap A) > 0$ which contradicts A being an atom. Hence $\alpha = \beta$. It follows that

$$P((-\infty, \alpha - \frac{1}{n}) \cap A) = 0, \quad P((\alpha + \frac{1}{n}, \infty) \cap A) = 0$$

and thus

$$P((\alpha - \frac{1}{n}, \alpha + \frac{1}{n}) \cap A) = P(A) > 0.$$

Let $n \to \infty$ to get $P(\{\alpha\} \cap A) = P(A) > 0$. So $P(\{\alpha\} \triangle A) = 0$.

(c) Let A and B be distinct atoms. Then $P(A \triangle B) > 0$ and therefore $P(AB \triangle \emptyset) = P(AB\emptyset^c) + P((AB)^c\emptyset) = P(AB)$, and we claim that

$$P(AB) = 0.$$

Since $AB \subset A$ and $AB \subset B$ and A and B are atoms we have

either $P(AB) = 0$, in which case the claim is true,

or else $P(AB) > 0$, in which case, since A and B are atoms, we have

$$P(B \backslash AB) = P(A \backslash AB) = 0,$$

which means that $P(A \triangle B) = 0$ which contradicts the assumption that $P(A \triangle B) > 0$.

(d) Let

$$\mathcal{A}_n = \{ \text{ distinct atoms which have probability at least } \frac{1}{n} \}.$$

There are at most n atoms in $\mathcal{A}_n$ since if $A_1, ..., A_m \in \mathcal{A}_n$, then

$$1 \geq P(\bigcup_{j=1}^{m} A_j) = \sum_{j=1}^{m} P(A_j) \geq \frac{m}{n},$$

which makes $m \leq n$ and $\text{card}(\mathcal{A}_n) \leq n$. So

$$\text{card}\{ \text{ all atoms } \} = \text{card}(\bigcup_n \mathcal{A}_n)$$

which is at most countable.

(e) A *partially ordered set* is a set S together with a relation, denoted $\leq$, on $S \times S$; that is, on pairs of elements of S. This relation satisfies

1. $x \leq x$,
2. $x \leq y$ and $y \leq x$ implies $x = y$,
3. $x \leq y$ and $y \leq z$ implies $x \leq z$.

A subset C of S is called a *chain* or a *totally ordered* subset, if every two elements of C are comparable; that is, if $x, y \in C$, then either $x \leq y$ or $y \leq x$. An *upper bound* of a set $A \subset S$ is an element y such that $x \leq y$ for all $x \in A$. A maximal element of S is any $y \in S$ satisfying $y \leq x$ implies $y = x$. *Zorn's lemma* says that if S is a partially ordered set in which every totally ordered subset has an upper bound, then S has a maximal element.

For a set $A \in \mathcal{B}$, recall

$$A^{\#} = \{B \in \mathcal{B} : P(A \triangle B) = 0\}.$$

Define the partial order on the equivalence class of sets to be $A^{\#} \leq B^{\#}$ iff there exists $A \in A^{\#}$, $B \in B^{\#}$ and $N \in \mathcal{B}$ such that $P(N) = 0$, and $A \subset B \cup N$. This is a well defined specification of the relation. If also $A' \in A^{\#}$ and $B' \in B^{\#}$, then

$$A'A^c = N_1, \quad A(A')^c = N_2,$$

where $P(N_i) = 0$, $i = 1, 2$ and

$$A = AA' + N_1, \quad A' = AA' + N_2.$$

Similarly,

$$B = BB' + N_3, \quad B' = BB' + N_4,$$

so

$$\begin{aligned} A' =& AA' + N_2 \subset B \bigcup N \bigcup N_2 \\ \subset& BB' \bigcup N_3 \bigcup N \bigcup N_2 \subset B' \bigcup N_5, \end{aligned}$$

where N_5 is an event with probability 0.

We have now defined a partial order relation on the equivalence classes of events since

1. $A^\# \leq A^\#$.
2. If $A^\# \leq B^\#$ and $B^\# \leq A^\#$ then

   $$A \subset B \bigcup N_1, \quad B \subset A \bigcup N_2$$

   and so

   $$A \subset B \bigcup N_1 \subset A \bigcup (N_1 \bigcup N_2)$$

   and therefore $P(A \triangle B) = 0$.
3. If $A^\# \leq B^\#$ and $B^\# \leq C^\#$ then $A \subset B \cup N_1$, $B \subset C \cup N_2$ and thus, $A \subset C \cup (N_1 \cup N_2)$ so $A^\# \leq C^\#$.

Now let $S^\# = \{A^\# : P^\#(A^\#) \leq a\}$. We claim that any subset $S_0^\#$ which is totally ordered has an upper bound. Write

$$S_0^\# = \{A_\alpha^\#, \alpha \in \Lambda\}$$

and set

$$p_\alpha = P^\#(A_\alpha^\#), \quad p_{S_0} = \sup_{\alpha \in \Lambda} P^\#(A_\alpha^\#).$$

By definition of supremum, there exists $\alpha_n \in \Lambda$ such that

$$p_{\alpha_n} = P^\#(A_{\alpha_n}^\#) \uparrow p_{S_0}.$$

To prove the claim, we consider two cases.

CASE 1. Suppose $p_\alpha < p_{S_0}$ for all $\alpha \in \Lambda$. Then we show for any $\alpha \in \Lambda$,

$$A_\alpha \leq \bigcup_n A_{\alpha_n}^\#$$

so $\bigcup_n A_{\alpha_n}^\#$ is an upper bound in $S^\#$. (Note, that it is relatively easy, by taking finite approximations to $\{\alpha_n\}$, to verify that the upper bound is, in fact, an element of $S^\#$.) To verify this, pick any $\alpha \in \Lambda$. Since $p_\alpha < p_{S_0}$, there exists α_n such that

$$p_\alpha < p_{\alpha_n}.$$

Then because of total ordering, either

$$A^{\#}_{\alpha} \le A^{\#}_{\alpha_n} \text{ or } A^{\#}_{\alpha_n} \le A_{\alpha}.$$

The latter is incompatible with $p_\alpha < p_{\alpha_n}$ and we conclude

$$A^{\#}_{\alpha} \le A^{\#}_{\alpha_n} \le \bigcup_n A^{\#}_{\alpha_n},$$

as needed.

CASE 2. Suppose there exists $\alpha^* \in \Lambda$ such that $p_{\alpha^*} = p_{S_0}$. Then we claim that $\cup_n A^{\#}_{\alpha_n} \cup A^{\#}_{\alpha^*}$ is an upper bound in $S^{\#}$. To see this, observe that for any α, either $p_\alpha < p_{S_0}$, in which case, as in Case 1,

$$A^{\#}_{\alpha} \subset \bigcup_n A^{\#}_{\alpha_n} \subset \bigcup_n A^{\#}_{\alpha_n} \bigcup A_{\alpha^*}$$

or, if $p_\alpha = p_{S_0}$, then either

$$A^{\#}_{\alpha} \le A^{\#}_{\alpha^*} \text{ or } A^{\#}_{\alpha^*} \le A^{\#}_{\alpha}.$$

In the first case, there exists $A_{\alpha^*} \in A^{\#}_{\alpha^*}$ and $A_\alpha \in A^{\#}_{\alpha}$ such that

$$A_{\alpha^*} \cup N \supset A_\alpha, \text{ and } P(A_{\alpha^*}) = P(A_\alpha).$$

So

$$P(A_\alpha A^c_{\alpha^*}) = P(A_\alpha) - P(A_\alpha A_{\alpha^*}) = 0,$$

and

$$P(A_{\alpha^*} A^c_{\alpha}) = P(A_{\alpha^*}) - P(A_\alpha A_{\alpha^*}) = 0,$$

so $P(A_{\alpha^*} \triangle A_\alpha) = 0$. Thus $A^{\#}_{\alpha^*} = A^{\#}_{\alpha}$ and

$$A^{\#}_{\alpha} \subset \bigcup A^{\#}_{\alpha_n} \bigcup A^{\#}_{\alpha^*}.$$

Consider the alternative case similarly.

By Zorn's lemma, there exists a maximal element $A^{\#}_{\max} \in S^{\#}$ such that $P^{\#}(A^{\#}_{\max}) \le a$. For the purposes of getting a contradiction, suppose $P^{\#}(A^{\#}_{\max}) < a$. We show this implies the existence of an atom.

Let

$$\mathcal{C}^{\#} = \{B^{\#} \neq \emptyset : B^{\#} \bigcap A^{\#}_{\max} = \emptyset\}.$$

For $B^{\#} \in \mathcal{C}^{\#}$, since $B^{\#} \cap A^{\#}_{\max} = \emptyset$,

$$P^{\#}(B^{\#}) + P^{\#}(A^{\#}_{\max}) = P^{\#}(B^{\#} \bigcup A^{\#}_{\max}) > a,$$

otherwise we would get a contradiction in the following way. If

$$P^{\#}(B^{\#}) + P^{\#}(A^{\#}_{\max}) = P^{\#}(B^{\#} \bigcup A^{\#}_{\max}) \le a,$$

then

$$B^{\#} + A^{\#}_{\text{max}} \in S^{\#}$$

so

$$A^{\#}_{\text{max}} \leq B^{\#} \cup A^{\#}_{\text{max}} \leq A^{\#}_{\text{max}}$$

where the last inequality follows by maximality. Thus from the previous line we would have

$$A^{\#}_{\text{max}} = B^{\#} \bigcup A^{\#}_{\text{max}}$$

and since $B^{\#} \neq \emptyset$ we get the desired contradiction to maximality of $A^{\#}_{\text{max}}$. Thus we conclude

$$P^{\#}(B^{\#}) > a - P^{\#}(A^{\#}_{\text{max}}) = \epsilon > 0.$$

Make a partial order on $\mathcal{C}^{\#}$ by defining $B^{\#} \leq C^{\#}$ iff there exists $B \in \mathcal{B}^{\#}$ and $C \in \mathcal{C}^{\#}$ such that $B \cup N \supset C$. Note that $\mathcal{C}^{\#}$ is ordered by the inverse relation to the one used for $S^{\#}$. As before, any totally ordered subset has an upper bound in $\mathcal{C}^{\#}$. The argument for this is similar to the argument used to show the corresponding fact for $S^{\#}$. For instance, if $S^{\#}_1$ is a totally ordered subset of $\mathcal{C}^{\#}$, write

$$S^{\#}_1 = \{B^{\#}_{\alpha}, \alpha \in \Lambda_1\}, \quad p_{\alpha} = P^{\#}(B^{\#}_{\alpha}) > \epsilon,$$

and define

$$p := \inf_{\alpha \in \Lambda_1} p_{\alpha} \geq \epsilon.$$

There exist $\alpha_n \in \Lambda_1$ such that $p_{\alpha_n} \downarrow p$. If for all $\alpha \in \Lambda_1$, $p_{\alpha} > p$, then $\cap_n B^{\#}_{\alpha_n}$ is the upper bound since for any $\alpha \in \Lambda_1$, there exist α_n such that $p_{\alpha_n} < p_{\alpha}$. Then we claim $B^{\#}_{\alpha} \leq \cap_k B^{\#}_{\alpha_k}$ since

$$P^{\#}(\bigcap_k B^{\#}_{\alpha_k}) \leq P^{\#}(B^{\#}_{\alpha_n}) < p_{\alpha} = P^{\#}(B^{\#}_{\alpha}).$$

Either $B^{\#}_{\alpha} \leq B^{\#}_{\alpha_n}$ or $B^{\#}_{\alpha_n} \leq B^{\#}_{\alpha}$ but the latter alternative is incompatible with the previous display so we get

$$B^{\#}_{\alpha} \leq B^{\#}_{\alpha_n} \leq \bigcap_n B^{\#}_{\alpha_n}.$$

Handling the case that some α satisfies $p_{\alpha} = p$ is similar to the procedure used in analyzing $S^{\#}$.

Again by Zorn's lemma, a maximal element $B^{\#}_{\text{max}} \in \mathcal{C}^{\#}$ exists. It follows that $B^{\#}_{\text{max}}$ is an atom. To see this, keep in mind $P^{\#}(B^{\#}_{\text{max}}) \geq \epsilon$. Let $B^{\#} \subset B^{\#}_{\text{max}}$, where $B^{\#}$ is the equivalence class of a set in $\mathcal{B}$. Then $P^{\#}(B^{\#}) \leq P^{\#}(B^{\#}_{\text{max}})$. If $B^{\#} \in \mathcal{C}^{\#}$, then $P^{\#}(B^{\#}_{\text{max}}) \leq P^{\#}(B^{\#})$ so we conclude that $P^{\#}(B^{\#}_{\text{max}}) = P^{\#}(B^{\#})$. Otherwise, if $B^{\#} \notin \mathcal{C}^{\#}$, then either $B^{\#} = \emptyset$ or

$B^{\#} \neq \emptyset$ and $B^{\#} \cap A^{\#}_{\max} \neq \emptyset$. This latter alternative is impossible since $B^{\#}_{\max} \in \mathcal{C}^{\#}$, so $B^{\#}_{\max} \cap A^{\#}_{\max} = \emptyset$, which implies $B^{\#} \cap A^{\#}_{\max} = \emptyset$.

Thus $B^{\#} \subset B^{\#}_{\max}$ implies $B^{\#} = \emptyset$ or $P^{\#}(B^{\#}) = P^{\#}(B^{\#}_{\max})$ and thus $B^{\#}_{\max}$ is an atom.

2.6.12. We show that iff $B \in \sigma(\mathcal{C})$, then there exists a countable family $\mathcal{C}_B \subset \mathcal{C}$ such that $B \in \sigma(\mathcal{C}_B)$.

To see this, we let

$$\mathcal{G} = \{B \subset \Omega : \exists \text{ a countable family } \mathcal{C}_B \subset \mathcal{C} \text{ such that } B \in \sigma(\mathcal{C}_B)\}.$$

Properties of $\mathcal{G}$:

(1) $\Omega \in \mathcal{G}$ since for any countable subset $\mathcal{C}' \subset \mathcal{C}$, we have $\Omega \in \sigma(\mathcal{C}')$.

(2) If $B \in \mathcal{G}$, then $B \in \sigma(\mathcal{C}_B)$ implies $B^c \in \sigma(\mathcal{C}_B)$. Hence $B^c \in \mathcal{G}$.

(3) If $B_n \in \mathcal{G}$ then $B_n \in \sigma(\mathcal{C}_{B_n}) \subset \sigma(\cup_n \mathcal{C}_{B_n})$, where $\mathcal{C}_{B_n}$ is a countable family and hence so is $\cup_n \mathcal{C}_{B_n}$. Therefore, $\bigcup_n B_n \in \sigma(\bigcup_n \mathcal{C}_{B_n})$ which implies $\bigcup_n B_n \in \mathcal{G}$. So $\mathcal{G}$ is a σ-field.

(4) $\mathcal{C} \subset \mathcal{G}$ since if $\Lambda \in \mathcal{C}$ then $\Lambda \in \sigma(\Lambda)$ and if we set $C_\Lambda = \{\Lambda\}$, then $\mathcal{C}_\Lambda$ is countable.

Thus $\mathcal{G} \supset \mathcal{C}$ which implies $\mathcal{G} \supset \sigma(\mathcal{C})$.

2.6.15. To check $\mathcal{S}_1\mathcal{S}_2 := \{S_1S_2 : S_i \in \mathcal{S}_i,\ i = 1, 2\}$ is a semi-algebra we must check three postulates.

1. $\emptyset \in \mathcal{S}_i$ for $i = 1, 2$ and therefore $\emptyset = \emptyset \cap \emptyset \in \mathcal{S}_1\mathcal{S}_2$. Similarly, we may prove $\Omega \in \mathcal{S}_1\mathcal{S}_2$.

2. If $S_1S_2 \in \mathcal{S}_1\mathcal{S}_2$ and $S_1'S_2' \in \mathcal{S}_1\mathcal{S}_2$ then
$$S_1S_2 \bigcap S_1'S_2' = S_1S_1' \bigcap S_2S_2' \in \mathcal{S}_1\mathcal{S}_2$$
since $S_1S_1' \in \mathcal{S}_1$ and $S_2S_2' \in \mathcal{S}_2$.

3. For $S_1 \in \mathcal{S}_i$, $i = 1, 2$ we have
$$\begin{aligned}(S_1S_2)^c &= S_1^c \bigcup S_2^c = S_1^cS_2 + S_1^cS_2^c + S_1S_2^c \\ &= \sum_{j=1}^{l} A_{1j}S_2 + \sum_{j,i} A_{1j}A_{2j} + \sum_{i=1}^{k} S_1A_{2i},\end{aligned}$$
where we assumed
$$S_1^c = \sum_{j=1}^{l} A_{1j}, \quad S_2^c = \sum_{i=1}^{k} A_{2i}.$$
This shows complements in $\mathcal{S}_1\mathcal{S}_2$ have the correct form.

To check that

$$\mathcal{A}(\mathcal{S}_1\mathcal{S}_2) = \mathcal{A}(\mathcal{S}_1 \bigcup \mathcal{S}_2),$$

note that the left side is of the form

$$\{\sum_{i=1}^{l} S_{1i}S_{2i},\ S_{1i} \in \mathcal{S}_1, S_{2i} \in \mathcal{S}_2\}.$$

Such sets as exhibited on the previous line are in $\mathcal{A}(\mathcal{S}_1 \bigcup \mathcal{S}_2)$ and therefore

$$\mathcal{A}(\mathcal{S}_1\mathcal{S}_2) \subset \mathcal{A}(\mathcal{S}_1 \bigcup \mathcal{S}_2).$$

Conversely

$$\mathcal{S}_1 \bigcup \mathcal{S}_2 \subset \mathcal{A}(\mathcal{S}_1\mathcal{S}_2)$$

and hence

$$\mathcal{A}(\mathcal{S}_1 \bigcup \mathcal{S}_2) \subset \mathcal{A}(\mathcal{S}_1\mathcal{S}_2).$$

2.6.16. Suppose $\{B_n\}$ are disjoint and $B_n \in \mathcal{B}$. Then $\sum_n B_n \in \mathcal{B}$ and $(\sum_n B_n)^c \in \mathcal{B}$ and therefore by assumption (c)

$$1 = Q(\sum_n B_n + (\sum_n B_n)^c) = \sum_n Q(B_n) + Q((\sum_n B_n)^c).$$

However, we also have, from finite additivity,

$$1 = Q(\sum_n B_n) + Q((\sum_n B_n)^c),$$

and therefore

$$Q(\sum_n B_n) = \sum_n Q(B_n).$$

2.6.17. We check $F_r^{\leftarrow}(y)$ is right continuous by showing that if $y_n \downarrow y$, then $F_r^{\leftarrow}(y_n) \downarrow F_r^{\leftarrow}(y)$. If this is not the case, then there exists L such that

$$F_r^{\leftarrow}(y_n) \downarrow L > F_r^{\leftarrow}(y).t$$

Suppose x is any value chosen so that

$$L > x > F_r^{\leftarrow}(y).$$

Then $F_r^{\leftarrow}(y_n) > x$ implies by the definition of $F_r^{\leftarrow}$ that $F(x) \leq y_n$ and therefore, by letting $n \to \infty$, that $F(x) \leq y$. On the other hand, since $x > F_r^{\leftarrow}(y)$, we have by definition that $F(x) > y$ so we conclude $F(x) = y$. This means $F_r^{\leftarrow}(y) \geq x$ which is a contridiction to the fact that $x > F_r^{\leftarrow}(y)$.

2.6.21. For a finitely additive measure μ satisfying $\mu(\Omega) = 1$, it need not be the case that $A_n \downarrow \emptyset$ implies $\mu(A_n) \downarrow 0$. Use Proposition 2.6.1. Let $\Omega = \{1, 2, \ldots, \}$ and let $\mathcal{A} = \{E \subset \Omega : E$ or E^c is finite. $\}$. Define

$$P(E) = \begin{cases} 0, & \text{if } E \text{ is finite,} \\ 1, & \text{if } E^c \text{ is finite.} \end{cases}$$

Then P is finitely additive.

Let $A_n = \{n, n+1, \ldots\} \in \mathcal{A}$ since A_n^c is finite. So $P(A_n) = 1$. Note $A_n \downarrow \emptyset$ but $1 = P(A_n) \not\to 0$.

2.6.23. Set

$$\mathcal{C} = \{(-\infty, \boldsymbol{x}] : \boldsymbol{x} \in \mathbb{R}\},$$

which is a π-system generating $\mathcal{B}(\mathbb{R}^d)$ and so if $P_1 = P_2$ on $\mathcal{C}$, then $P_1 = P_2$ on $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R}^d)$.

Chapter 3 Solutions

3.4.1. If $1_A \in \mathcal{B}$, then
$$A = 1_A^{-1}\{1\} \in \mathcal{B}.$$
Conversely, suppose $A \in \mathcal{B}$. Then if I' is an interval,
$$1_A^{-1}(I') = \{\omega : 1_A(\omega) \in I'\}.$$
Consider the following cases.

1. If $I' \supset [0,1]$, then $1_A^{-1}(I') = \Omega \in \mathcal{B}$.
2. If $0 \in I'$ but $1 \notin I'$ then $1_A^{-1}(I') = A^c \in \mathcal{B}$.
3. If $1 \in I'$ but $0 \notin I'$ then $1_A^{-1}(I') = A \in \mathcal{B}$.
4. If I' contains neither 0 nor 1 then
$$1_A^{-1}(I') = \emptyset \in \mathcal{B}.$$

This suffices to show 1_A is measurable with respect to $\mathcal{B}$ by Proposition 3.2.1.

3.4.2. We have
$$\begin{aligned}
\sigma(X_1) =& \{\emptyset, \Omega\} \\
\sigma(X_2) =& \sigma(1_{\{1/2\}}) = \{\emptyset, \Omega, \{\frac{1}{2}\}, \{\frac{1}{2}\}^c\} \\
\sigma(X_3) =& \{\emptyset, \Omega, \mathbb{Q}, \mathbb{Q}^c\}
\end{aligned}$$

3.4.4. If $X \in \mathcal{B}/\mathcal{B}(\mathbb{R})$, then since $\{x\} \in \mathcal{B}(\mathbb{R})$,
$$X^{-1}(\{x\}) \in \mathcal{B},$$
for all $x \in \mathcal{B}(\mathbb{R})$.

Conversely, suppose for all $x \in \mathcal{R}$ that $X^{-1}(\{x\}) \in \mathcal{B}$. Then for any $B \in \mathcal{B}(\mathbb{R})$,
$$\begin{aligned}
X^{-1}(B) =& \{\omega : X(\omega) \in B\} = \{\omega : X(\omega) \in B \bigcap \mathcal{R}\} \\
=& \bigcup_{r \in \mathcal{R}, r \in B} \{w : X(\omega) = r\} \\
=& \bigcup_{r \in \mathcal{R}, r \in B} X^{-1}(\{r\}) \in \mathcal{B}.
\end{aligned}$$

3.4.5. The variable $Y = F(X)$ is measurable by composition:
$$\begin{aligned}
X &: (\Omega, \mathcal{B}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\
F &: (\mathbb{R}, \mathcal{B}(\mathbb{R})) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))
\end{aligned}$$

since F is monotone and hence measureable.

Since F is continuous, $P[X \leq x] = P[X < x]$. So from the properties of $F^{\leftarrow}$ we have

$$P[F(X) \geq y] = P[X \geq F^{\leftarrow}(y)] = 1 - F(F^{\leftarrow}(y)).$$

But

$$F(\inf\{u : F(u) \geq y\}) = y$$

when F is continuous.

3.4.8. We write

$$Z = X1_A + Y1_{A^c}.$$

If $A \in \mathcal{B}$ then 1_A and 1_{A^c} are both random variables by 3.4.1. Products of random variables are random variables and sums of random variables are random variables. This suffices.

3.4.11. X_t is a random variable since for fixed t, $X_t = 1_{\{t\}}$ and $\{t\}$ is measurable. So

$$\sigma(X_t) = \{\emptyset, \Omega, \{t\}, \{t\}^c\}.$$

We claim

$$\begin{aligned} LHS = & \bigvee_{t \in [0,1]} \sigma(X_t) \\ =& \{A \subset [0,1] : A \text{ is countable or } A^c \text{ is countable.}\} = RHS. \end{aligned}$$

Let

$$\mathcal{C} = \{\emptyset, \Omega, \{t\}, \{t\}^c; t \in [0,1]\}$$

so that

$$LHS = \sigma(\mathcal{C}) = \bigvee_{t \in [0,1]} \sigma(X_t).$$

Clearly $RHS \subset LHS$ since the LHS contains one point sets and is closed under countable union.

Likewise, $A \in \mathcal{C}$ implies that $A \in RHS$ so that $LHS \subset RHS$.

3.4.12. To show that monotone f is measurable $\mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$ it suffices to show that

$$\{u : f(u) \leq x\} = \mathcal{B}(\mathbb{R}).$$

However, by monotonicity, the left side is a semi-infinite interval which is certainly a Borel set. From Proposition 3.2.1 or Corollary 3.2.1, this suffices.

3.4.14. The function f is usc iff $\{t : f(x) < \alpha\} = f^{-1}(-\infty, \alpha)$ is *open* in $\mathbb{R}$. If $\mathcal{C} = \{(-\infty, \lambda) : \lambda \in \mathbb{R}\}$, then $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$ and $f^{-1}(\mathcal{C}) \subset \mathcal{B}(\mathbb{R})$. This means $f \in \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$.

3.4.15. Let $f(x) = 1_{(a,b]}$ and define

$$f_n(x) = \begin{cases} 1, & \text{if } a + \frac{1}{n} < x \le b, \\ 0, & \text{if } x \le a, \text{ or } x \ge b + \frac{1}{n}, \\ \text{linear}, & \text{otherwise.} \end{cases}$$

For $x \in (a,b]$, $f(x) = 1$ and $f_n(x) = 1$, provided n is so large that $a + \frac{1}{n} < x$. For $x \le a$ we have $f_n(x) = f(x) = 0$ while for $x > b$ we have $f(x) = f_n(x) = 0$ provided $b + \frac{1}{n} < x$.

3.4.16. If $\mathcal{B}(\mathbb{R}) \subset \mathcal{B}$ and f is continuous, then f is measurable so

$$f^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{B}(\mathbb{R}) \subset \mathcal{B}$$

and therefore $f \in \mathcal{B}$.

Conversely, suppose for any continuous function f that we have $f \in \mathcal{B}$. This implies by definition that

$$f^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{B}.$$

This says that for any $\Lambda \in \mathcal{B}(\mathbb{R})$

$$\{x : f(x) \in \Lambda\} \in \mathcal{B}$$

for any continuous function f. Let $f(x) = x$ and the previous display reads $\Lambda \in \mathcal{B}$. This means $\mathcal{B}(\mathbb{R}) \subset \mathcal{B}$.

Now let $\mathcal{F} = \sigma(f, f \in C(\mathbb{R}))$ be the smallest σ-field containing all continuous functions on $\mathbb{R}$. From the previous discussion, we get

$$\mathcal{B}(\mathbb{R}) \subset \mathcal{F}.$$

But if f is continuous, f is $\mathcal{B}(\mathbb{R})$ measurable. So if any continuous function is $\mathcal{B}(\mathbb{R})$ measurable, the smallest σ-field generated by the continuous functions must be contained in $\mathcal{B}(\mathbb{R})$. Hence $\mathcal{F} = \mathcal{B}(\mathbb{R})$.

3.4.17. Start by assuming $T \in \mathcal{B}/\mathcal{B}'$. Then for $B' \in \mathcal{B}'$, we have

$$T_n^{-1}(B') = \{\omega \in A_n : T_n(\omega) \in B'\} = A_n \cap T^{-1}(B') \in \mathcal{B}_n.$$

Thus $T_n^{-1}(\mathcal{B}') \subset \mathcal{B}_n$.

Conversely, suppose for each n that $T_n \in \mathcal{B}_n/\mathcal{B}'$. Then for $B' \in \mathcal{B}'$ we have

$$\begin{aligned} T^{-1}(B') =& \{\omega : T(\omega) \in B') = \bigcup_n \{\omega \in A_n : T(\omega) \in B'\} \\ =& \bigcup_n \{\omega \in A_n : T_n(\omega) \in B'\} = \bigcup_n T_n^{-1}(B'). \end{aligned}$$

Since $T_n^{-1}(B') \in \mathcal{B}_n \subset \mathcal{B}$, we have $T^{-1}(B') \in \mathcal{B}$.

3.4.19. Suppose first that $X = Y \circ T$ and we show $X \in \sigma(T)$. For any $A \in \mathcal{B}(\mathbb{R})$, we need to show

$$X^{-1}(A) \in \sigma(T) = \{T^{-1}(B_2) : B_2 \in \mathcal{B}_2\}.$$

This follows from

$$X^{-1}(A) = T^{-1}(Y^{-1}(A))$$

since $Y^{-1}(A) \in \mathcal{B}_2$.

Conversely, suppose $X \in \sigma(T)$ which means that for all $A \in \mathcal{B}(\mathbb{R})$,

$$X^{-1}(A) \in \{T^{-1}(B_2) : B_2 \in \mathcal{B}_2\}.$$

Suppose, for simplicity, that $X \geq 0$, since otherwise we would just split X into positive and negative parts. Then we write

$$X = \lim_{n\to\infty} \uparrow \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{\left[X \in [\frac{k-1}{2^n}, \frac{k}{2^n})\right]} + n 1_{[X \geq n]}$$

and for some sets $B_{kn}, B_n \in \mathcal{B}_2$ the above equals

$$= \lim_{n\to\infty} \uparrow \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{T^{-1}(B_{kn})} + n 1_{T^{-1}(B_n)}$$

and thus for any $\omega_1 \in \Omega_1$

$$X(\omega_1) = \lim_{n\to\infty} \uparrow \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{B_{kn}}(T\omega_1) + n 1_{B_n}(T\omega_1).$$

Define

$$Y_n(\omega_2) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{B_{kn}}(\omega_2) + n 1_{B_n}(\omega_2),$$

and

$$Y = \limsup_{n\to\infty} Y_n.$$

Then $X = Y \circ T$ as required.

Chapter 4 Solutions

4.6.2. We use the following useful notation. If A is any set, define

$$A^{(0)} = A^c, \quad A^{(1)} = A.$$

We need the fact that if $B_1, \dots, B_n$ are independent events, so are $B^{(\epsilon_1)}, \dots, B^{(\epsilon_n)}$ for any choice of $\boldsymbol{\epsilon} := (\epsilon_1, \dots, \epsilon_n) \in \{0,1\}^n$.

Suppose $B_1, \dots, B_n$ are independent subsets of a space Ω_n satisfying

$$1 > P(B_i^{(\epsilon_i)}) > 0; \ \epsilon_i \in \{0,1\}, \ i = 1, \dots, n.$$

Then

$$P(\bigcap_{i=1}^{n} B_i^{(\epsilon_i)}) = \prod_{i=1}^{n} P(B_i^{(\epsilon_i)}) > 0,$$

implies $\bigcap_{i=1}^{n} B_i^{(\epsilon_i)} \neq \emptyset$. The sets

$$\{\bigcap_{i=1}^{n} B_i^{(\epsilon_i)} : (\epsilon_1, \dots, \epsilon_n) \in \{0,1\}^n\}$$

partition Ω_n. So if $|A|$ is the cardinality of A,

$$|\Omega_n| = \sum_{\boldsymbol{\epsilon} \in \{0,1\}^n} |\bigcap_{i=1}^{n} B_i^{(\epsilon_i)}| \geq |\{0,1\}| = 2^n,$$

since $\bigcap_{i=1}^{n} B_i^{(\epsilon_i)} \neq \emptyset$ and hence must contain at least one sample point. So having n independent events, requires the space to have at least 2^n sample points.

It is easy to see that 2^n is really the correct minimum number. Let

$$\Omega_n = \{0,1\}^n, \ P((\epsilon_1, \dots, \epsilon_n)) = \frac{1}{2^n}$$

for all $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n) \in \Omega_n$ and set

$$B_i = \{\boldsymbol{\epsilon} : \epsilon_i = 1\}.$$

Then $B_1, \dots, B_n$ are independent and $1 > P(B_i^{(\epsilon_i)}) > 0$. Since $|\Omega_n| = 2^n$, we conclude the sample space cannot contain fewer than 2^n points if n independent events exist.

4.6.5. (a) If X is independent of itself, $\mathcal{B}(X)$ is almost trivial. Therefore, since $X \in \mathcal{B}(X)$, we have that there exists $c \in \mathbb{R}$ such that

$$P[X = c] = 1.$$

(b) If X is independent of $g(X)$, then $g(X)$ is independent of $g(X)$ and hence by (a), there is some c such that $P[g(X) = c] = 1$.

4.6.6. We have on the one hand that

$$P[X_1 \le x_1, \dots, X_n \le x_n] = \prod_{i=1}^{n} F(x_i),$$

for all $x_i \in \mathbb{R}$, $i = 1, \dots, n$. On the other hand,

$$P[X_{\pi(1)} \le x_1, \dots, X_{\pi(n)} \le x_n] = \prod_{i=1}^{n} F(x_i),$$

so

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

4.6.11. Pick c_n to satisfy

$$P[|X_n| > c_n/n] \le \frac{1}{2^n},$$

so that

$$\sum_n P\left[|\frac{X_n}{c_n}| > \frac{1}{n}\right] < \infty.$$

By the Borel-Cantelli lemma, $P\left[|\frac{X_n}{c_n} > \frac{1}{n} \text{ i.o.}\right] = 0$, and therefore

$$1 = P\{\liminf_{n\to\infty}[|\frac{X_n}{c_n}| \le \frac{1}{n}]\} \le P\{\lim_{n\to\infty}|\frac{X_n}{c_n}| = 0\}.$$

4.6.12. Given any $\epsilon > 0$, $a_n/b_n \to 1$ means

$$b_n(1-\epsilon) \le a_n \le b_n(1+\epsilon),$$

for all large n, say $n \ge n_0 = n_0(\epsilon)$. Therefore

$$\sum_{n \ge n_0} b_n(1-\epsilon) \le \sum_{n \ge n_0} a_n \le \sum_{n \ge n_0} b_n(1+\epsilon)$$

and the result follows.

4.6.13. Let

$$(B_1, B_2, \dots) = (A_1, A_4, A_7, \dots)$$

so that $\{B_n\}$ are independent events. Also

$$P(B_n) = P(A_1) = p^2 q$$

so that $\sum P(B_n) = \sum_n p^2 q = \infty$ and by the Borel 0 - 1 law we have

$$P(B_n \text{ i.o.}) = 1.$$

Since

$$\limsup_{n\to\infty} B_n \subset \limsup_{n\to\infty} A_n$$

we have

$$1 = P(\limsup_{n\to\infty} B_n) \leq P(\limsup_{n\to\infty} A_n).$$

4.6.14. For $n \geq 1$, define the events

$$A_n = \bigcup_{j=0}^{2^n-n} [X_{2^n+j} = 1, \dots, X_{2^n+j+n-1} = 1, X_{2^n+j+n} = 0].$$

Taking complements we have

$$A_n^c = \bigcap_{j=0}^{2^n-n} [X_{2^n+j} = 1, \dots, X_{2^n+j+n-1} = 1, X_{2^n+j+n} = 0]^c$$

and retaining only certain terms in the intersection gives

$$\begin{aligned} A_n^c \subset &[X_{2^n+0} = 1, \dots, X_{2^n+n-1} = 1, X_{2^n+n} = 0]^c \\ &\bigcap [X_{2^n+n+1} = 1, \dots, X_{2^n+2n} = 1, X_{2^n+2n+1} = 0]^c \\ &\bigcap [X_{2^n+2n+2} = 1, \dots, X_{2^n+3n+1} = 1, X_{2^n+3n+2} = 0] \\ &\bigcap \dots, \end{aligned}$$

which is the intersection of events depending on disjoint blocks of X's and which are therefore, by the groupings lemma, independent. So $P(A_n^c) \leq \prod(1 - p^n q)$ where the number of terms in the product is the number of disjoint blocks of length $n+1$ which can be crammed in the interval $[2^n, 2^{n+1}]$. This is about

$$\frac{2^{n+1} - 2^n}{n+1} = \frac{2^n(2-1)}{n+1} = \frac{2^n}{n+1}.$$

Therefore,

$$P(A_n^c) \leq (1 - p^n q)^{2^n/n+1}$$

and using the inequality $1 - x \leq e^{-x}$ for $0 \leq x \leq 1$

$$P(A_n^c) \leq \exp\{-qp^n \frac{2^n}{n+1}\} = \exp\{-q\frac{(2p)^n}{n+1}\}$$

which leads to

$$P(A_n) = 1 - P(A_n^c) \geq 1 - \exp\{-q\frac{(2p)^n}{n+1}\}.$$

Now $\{A_n\}$ are independent, so to prove $P(A_n \text{ i.o.}) = 1$ we must show $\sum_n P(A_n) = \infty$. For $p = \frac{1}{2}$

$$\sum_n P(A_n) \geq \sum_n (1 - \exp\{-\frac{1/2}{n+1}\}) = \infty$$

since

$$1 - e^{-\frac{1/2}{n+1}} \sim \frac{1/2}{n+1}, \quad n \to \infty$$

which is not summable. For $p > \frac{1}{2}$ we also have $2p > 1$ and $\frac{(2p)^n}{n+1} \to \infty$ so $\sum_n P(A_n) = \infty$.

4.6.15. If $E(\prod_{i \in I} Y_i) = \prod_{i \in I} E(Y_i)$ whenever $Y_i \in \mathcal{B}_i$, then for any $A_i \in \mathcal{B}_i$, take $Y_i = 1_{A_i}$ and

$$E(\prod_{i \in I} Y_i) = P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i) = \prod_{i \in I} E(Y_i)$$

and so $\{A_i, i \in I\}$ are independent and $\{\mathcal{B}_i, i \in I\}$ are independent σ-fields.

Conversely, suppose $\{\mathcal{B}_i, i \in I\}$ are independent σ-fields. For $A_i \in \mathcal{B}_i$ and $Y_i = 1_{A_i}$ we have

$$E(\prod_{i \in I} Y_i) = \prod_{i \in I} E(Y_i).$$

Next, if $\{Y_i, i \in I\}$ are simple, then suppose

$$Y_i = \sum_j x_{i,j} 1_{A_{i,j}}, \quad A_{i,j} \in \mathcal{B}_i.$$

We then have

$$\begin{aligned}
E(\prod_{i \in I} Y_i) &= E\Big(\prod_{i \in I} \sum_{j(i)} x_{i,j(i)} 1_{A_{i,j(i)}}\Big) \\
&= E\Big(\sum_{j(i), i \in I} \prod_{i \in I} x_{i,j(i)} 1_{\bigcap_{i \in I} A_{i,j(i)}}\Big) \\
&= \sum_{j(i), i \in I} \prod_{i \in I} x_{i,j(i)} P\Big(\bigcap_{i \in I} A_{i,j(i)}\Big) \\
&= \sum_{j(i), i \in I} \prod_{i \in I} [x_{i,j(i)} P(A_{i,j(i)})] \\
&= \prod_{i \in I} \sum_{j(i)} x_{i,j(i)} P(A_{i,j(i)}) \\
&= \prod_{i \in I} E(Y_i).
\end{aligned}$$

(If the notation makes following this difficult, write out the argument assuming that $I = \{1, 2\}$.)

Finally, if Y_i is a general, non-negative $\mathcal{B}_i$-measurable function, there exists $Y_i^{(n)} \in \mathcal{B}_i$, such that $Y_i^{(n)}$ is simple and $0 \leq Y_i^{(n)} \uparrow Y_i$. It then follows that

$$\prod_{i \in I} Y_i^{(n)} \uparrow \prod_{i \in I} Y_i$$

and by the monotone convergence theorem

$$E(\prod_{i \in I} Y_i^{(n)}) \uparrow E(\prod_{i \in I} Y_i)$$

and from the previous step, the left side is

$$\prod_{i \in I} E(Y_i^{(n)}) \uparrow \prod_{i \in I} E(Y_i),$$

again using the monotone convergence theorem.

(b) If for example $\mathcal{B}_1$ is independent of $\mathcal{B}_2$ and $\mathcal{B}_i \supset \mathcal{B}_i'$ for $i = 1, 2$, then $\mathcal{B}_i'$, $i = 1, 2$ are independent since if $A_i \in \mathcal{B}_i'$ then $A_i \in \mathcal{B}_i$ and hence A_1, A_2 are independent. So if $X_t, t \in T$ are independent, then by definition $\sigma(X_t), t \in T$ are independent and since $\sigma(f(X_t)) \subset \sigma(X_t)$ the result follows.

4.6.16. Kolmogorov's 0-1 law implies that $P[X_n \text{ converges}] = 0$ or 1. If $P[X_n \text{ converges}] = 1$, then there exists $c \in [-\infty, \infty]$ such that $P[X_n \to c] = 1$, since $\lim_{n \to \infty} X_n$ is a tail random variable of an independent sequence and is hence almost surely constant. Suppose $|c| < \infty$. (Modest changes are necessary if $c = \pm\infty$.) Then for any $\varepsilon > 0$,

$$P[X_n \in (c - \varepsilon, c + \varepsilon)^c \text{ i.o.}] = 0,$$

so by Borel's 0–1 law,

$$\infty > \sum_n P[X_n \in (c - \varepsilon, c + \varepsilon)^c].$$

Since $\{X_n\}$ and iid sequence, we have

$$P[X_n \in (c - \varepsilon, c + \varepsilon)^c] = 0$$

(otherwise, the sum would diverge since the sum consists of equal terms by the iid assumption) so

$$P[X_1 \in (c - \varepsilon, c + \varepsilon)] = 1.$$

This is true for any $\varepsilon > 0$ so let $\varepsilon \downarrow 0$ to get $P[X = c] = 1$. This contradicts the assumption that the sequence does not consist of constants with probability one.

4.6.17. (b) If N is a $N(0,1)$ random variable, we have

$$\begin{aligned}P[|N|>(1\pm\varepsilon)\sqrt{2\log n}] =&2P[N>(1\pm\varepsilon)\sqrt{2\log n}]\\ \sim&2\frac{\exp\{-((1\pm\varepsilon)^2 2\log n/2\}}{(1\pm\varepsilon)\sqrt{2\log n}}\\ =&\frac{c}{n^{(1\pm\varepsilon)^2}\sqrt{\log n}}.\end{aligned}$$

Therefore

$$\sum_n P[|N|>(1+\varepsilon)\sqrt{2\log n}]<\infty,\quad \sum_n P[|N|>(1-\varepsilon)\sqrt{2\log n}]=\infty,$$

and

$$P[\frac{|X_n|}{\sqrt{\log n}}>(1+\varepsilon)\sqrt{2}\text{ i.o.}]=0,\quad P[\frac{|X_n|}{\sqrt{\log n}}>(1-\varepsilon)\sqrt{2}\text{ i.o. }]=1.$$

Thus

$$P[\limsup_{n\to\infty}\frac{|X_n|}{\sqrt{\log n}}=\sqrt{2}]=1.$$

(c) Let X have a Poisson distribution with parameter λ. Since

$$P[X\geq n]=\sum_{j=n}^{\infty}e^{-\lambda}\lambda^j/j!\geq e^{-\lambda}\lambda^n/n!,$$

we merely have to prove the upper bound.

We use the relation between the Poisson distribution and exponential distribution. Let $\{E_n, n\geq 1\}$ be iid unit exponential random variables so that $\{E_n/\lambda, n\geq 1\}$ are iid exponential random variables with parameter lambda. In the time interval $[0,1]$, a Poisson process of rate λ has at least n points iff the time of the nth occurrence is before time 1. Therefore

$$\begin{aligned}P[X\geq n] =&P[\sum_{i=1}^{n}E_i/\lambda\leq 1]=P[\sum_{i=1}^{n}E_i\leq\lambda]\\ =&\int_0^{\lambda}e^{-u}\frac{u^{n-1}}{(n-1)!}du\end{aligned}$$

and since $e^{-u}\leq 1$ for $u>0$, we get an upper bound

$$\leq 1\cdot\int_0^{\lambda}\frac{u^{n-1}}{(n-1)!}du=\frac{\lambda^n}{n!},$$

as needed to be shown.

We show now that

$$P[\limsup_{n\to\infty} \frac{X_n}{\log n/\log_2 n} = 1] = 1,$$

where $\log_2 n = \log(\log n)$. It suffices to show

$$\sum_n P[X_n \geq \alpha(\log n/\log_2 n)] \begin{cases} < \infty, & \text{if } \alpha > 1, \\ = \infty, & \text{if } \alpha < 1. \end{cases}$$

Set $m(n) = [\alpha \log n/\log_2 n]$ and note

$$P[X \geq m(n)] \leq \frac{\lambda^{m(n)}}{m(n)!}$$

and applying Stirling's formula to the denominator, this expression is asymptotic to

$$\begin{aligned} &\sim c \frac{\lambda^{m(n)}}{e^{-m(n)} m(n)^{m(n)+1/2}} = c\Big(\frac{\lambda e}{m(n)}\Big)^{m(n)} \frac{1}{\sqrt{m(n)}} \\ &= c \frac{\exp\{-m(n)\log\big(m(n)/\lambda\big)\}}{\sqrt{m(n)}} \\ &= c \frac{\exp\{-\alpha \frac{\log n}{\log_2 n} \cdot \log\big(\frac{\log n}{\lambda \log_2 n}\big)\}}{\sqrt{\alpha \frac{\log n}{\log_2 n}}} \\ &= c \exp\{-\alpha \log n (1 - \frac{\log(\lambda \log_2 n)}{\log_2 n})\} \frac{1}{\sqrt{\alpha \frac{\log n}{\log_2 n}}}. \end{aligned}$$

Since $\alpha > 1$, we may find $\alpha > \alpha' > 1$ and n so large that

$$\alpha[1 - \log(\lambda \log_2 n)/\log_2] \geq \alpha' > 1.$$

Then we get an upper bound for the tail probability as follows:

$$\leq \frac{\exp\{-\alpha' \log n\}}{\sqrt{\alpha \frac{\log n}{\log_2 n}}} = \frac{1}{n^{\alpha'} \sqrt{\alpha \frac{\log n}{\log_2 n}}}$$

When $\alpha > \alpha' > 1$, this is summable.

Similarly, when $\alpha < 1$, one gets the probability sum appearing in the Borel 0-1 Law to diverge.

4.6.18. Set $\mathcal{C}_1 = \{A\}$, and $\mathcal{C}_2 = \mathcal{P}$ and it follows that $\mathcal{C}_1$ is independent of $\mathcal{C}_2$. This implies that $\mathcal{B}(\mathcal{C}_1)$ is independent of $\mathcal{B}(\mathcal{C}_2)$. Therefore, A is independent of A and $P(A) = 0$ or 1.

4.6.19. Let $\Omega = \{0,1\}^2$. Define $X_1(i,j) = i$, $X_2(i,j) = j$ for $(i,j) \in \{0,1\}^2$. Now define P_1 and P_2 by

	(0,0)	(0,1)	(1,0)	(1,1)
P_1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
P_2	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

Then under P_1, we have X_1 independent of X_2 but not under P_2 since

$$P_2[X_1 = 0, X_2 = 0] = P_2(\{0,0\}) = \frac{1}{2}$$

and

$$\begin{aligned} P_2[X_1 = 0]P_2[X_2 = 0] =& P_2\Big(\{(0,0),(0,1)\}\Big)P_2\Big(\{(1,0),(0,0)\}\Big) \\ =& (\frac{1}{2} + \frac{1}{8})(\frac{1}{8} + \frac{1}{2}) = \frac{5}{8} \cdot \frac{5}{8} = \frac{25}{64} \neq \frac{1}{2}. \end{aligned}$$

4.6.20. (c) Note that

$$P[X_1X_2 = 1] = P[X_1 = 1, X_2 = 1 \text{ or } X_1 = -1, X_2 = -1] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

and similarly $P[X_1X_2 = -1] = 1/2$. Then X_i is independent of X_1X_2, for $i = 1,2$. To see this, note

$$\begin{aligned} P[X_1X_2 = 1, X_1 = 1] =& P[X_1 = 1, X_2 = 1] = \frac{1}{4} \\ =& P[X_1X_2 = 1]P[X_1 = 1] = (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4} \end{aligned}$$

and similarly for other possible values for the mass function. However,

$$P[X_1 = 1, X_2 = 1, X_1X_2 = -1] = 0 \neq P[X_1 = 1]P[X_2 = 1]P[X_1X_2 = 1]$$

and thus X_1, X_2, X_1X_2 are not independent.

4.6.22. Pick any $J \subset \{1,2,\dots\}$ and define

$$B_n = \begin{cases} A_n, & \text{if } n \in J, \\ A_n^c, & \text{if } n \in J^c. \end{cases}$$

Then $\{B_n\}$ are independent and $\sum_n P(B_n) = \infty$ so we conclude from the Borel 0–1 Law that

$$P(B_n \text{ i.o.}) = 1.$$

Taking complements, this also means that

$$0 = P(\liminf_{n\to\infty} B_n^c) = \lim_{n\to\infty} \uparrow P(\bigcap_{k\geq n} B_k^c).$$

Interchange the roles of J and J^c and we may conclude

$$0 \leq \lim_{n\to\infty} \uparrow P(\bigcap_{k\geq n} B_k) = 0.$$

Therefore, for all n,

$$P(\bigcap_{k\geq n} B_k) = 0.$$

For the purposes of getting a contradiction, suppose B is an atom so that $P(B) > 0$. Define J in the following particular way:

$$J := \{n \geq 1 : P(A_n B) \geq P(A_n^c B)\}$$

and as above, set

$$B_n = \begin{cases} A_n, & \text{if } n \in J, \\ A_n^c, & \text{if } n \in J^c. \end{cases}$$

Note with this definition, that $P(B_n B) \geq P(B_n^c B)$. Therefore, since

$$P(B) = P(B_n B) + P(B_n^c B) \leq 2P(B_n B),$$

we conclude

$$0 < \frac{1}{2}P(B) \leq P(B_n B).$$

Now, $B_n B \subset B$ and since B is assumed to be an atom, either $P(B_n B) = 0$ (which cannot be the case since we know $P(B_n B) > \frac{1}{2}P(B) > 0$) or else

$$P(B_n B) = P(B).$$

This last equality holds for all n. Since $P(B) > 0$ we have for the conditional probability measure

$$P(B_n | B) = 1$$

for all n and hence for any n

$$P(\bigcap_{k\geq n} B_k | B) = 1.$$

Therefore, for any n,

$$P(\bigcap_{k\geq n} B_k B) = P(B) > 0.$$

(b) We have

$$P[l_n = k] = P[d_n = 0, \ldots, d_{n+k-1} = 0, d_{n+k} = 1] = 2^{-(k+1)}.$$

Note

$$[l_n \geq r] = P[d_n = 0, \ldots, d_{n+r-1} = 0] = 2^{-r}$$

which also shows that

$$[l_n \geq r] \in \sigma(d_n, \ldots, d_{n+r-1}).$$

(c) This follows from $[l_n = 0] = [d_n = 0]$ and the fact that $\{d_n\}$ is iid.
(d) From the Borel 0-1 law, $P\{[l_n = 0] \text{ i.o. }\} = 1$ iff

$$\infty = \sum P[l_n = 0] = \sum_n \frac{1}{2}.$$

(e) Note

$$[l_{2n} = 1] = [d_{2n} = 0, d_{2n+1} = 1] \in \sigma(d_{2n}, d_{2n+1})$$

and by the groupings lemma, the sigma-fields $\sigma(d_{2n}, d_{2n+1})$ are independent for different n. Since

$$\sum_n P[l_{2n} = 1] = \sum_n \left(\frac{1}{2}\right)^2 = \infty,$$

$P\{[l_{2n} = 1] \text{ i.o. }\} = 1$ and hence $P\{[l_n = 1] \text{ i.o. }\} = 1$.
(f) We have

$$\sum_n P[l_n \geq (1+\epsilon)\log_2 n] \leq \sum_n P[l_n \geq [(1+\epsilon)\log_2 n]]$$
$$= \sum_n \left(\frac{1}{2}\right)^{[(1+\epsilon)\log_2 n]} = \sum_n \left(\frac{1}{2}\right)^{(1+\epsilon)\log_2 n + \theta(n)}$$

where $|\theta(n)| \leq 1$. This series converges or diverges according to whether $\sum_n n^{-(1+\epsilon)}$ converges or diverges but this series converges.
Therefore, by the Borel-Cantelli lemma, for any $\epsilon > 0$,

$$P\{[\frac{l_n}{\log_2 n} \geq 1 + \epsilon] \text{ i.o. }\} = 0.$$

This means for any $\epsilon_k \downarrow 0$

$$P[\limsup_{n\to\infty} \frac{l_n}{\log_2 n} \leq 1 + \epsilon_k] = 1$$

and intersecting over k

$$P\{\bigcap_k [\limsup_{n\to\infty} \frac{l_n}{\log_2 n} \le 1 + \epsilon_k]\} = P[\limsup_{n\to\infty} \frac{l_n}{\log_2 n} \le 1] = 1.$$

(g) Let $r(n) \uparrow \infty$ be a sequence of integers. In particular, we will let $r(n) = [\log_2 n]$. Define $n(1) = 1$, and $n(k+1) = n(k) + r(n(k))$. Then since

$$[l_n \ge r] \in \sigma(l_n, l_{n+1}, \ldots, l_{n+r-1})$$

we have from the groupings lemma that

$$[l_{n(k)} \ge r(n(k))] \in \sigma(d_{n(k)}, \ldots, d_{n(k+1)-1}),$$

and therefore these events are independent. If we show

$$P\{[l_{n(k)} \ge r(n(k))] \text{ i.o. }\} = 1$$

then it will also follow that

$$P\{[l_n \ge r(n)] \text{ i.o. }\} = 1.$$

It suffices to show

$$\sum_k P[l_{n(k)} \ge r(n(k))] = \infty.$$

Since $n(k+1) - n(k) = r(n(k))$, we have

$$\sum_k P[l_{n(k)} \ge r(n(k))] = \sum_k 2^{-r(n(k))} = \sum_k 2^{-r(n(k))} \frac{n(k+1) - n(k)}{r(n(k))}$$

and because $r(n)$ is non-decreasing, this is bounded below by

$$\begin{aligned} &\ge \sum_k \sum_{n=n(k)}^{n(k+1)-1} \frac{2^{-r(n)}}{r(n)} \\ &= \sum_k \frac{2^{-r(n)}}{r(n)} = \sum_n \frac{1}{n \log_2 n} = \infty. \end{aligned}$$

CHAPTER 5 SOLUTIONS

5.5. (b) We use Fubini's theorem. Let $A = \{a_1, a_2, \dots\}$ be the countable set of atoms of F. Write

$$\begin{aligned}
E(F(X)) &= \int_{\mathbb{R}} F(x)F(dx) = \int_{\mathbb{R}} \left(\int_{y \le x} F(dy)\right) F(dx) \\
&= \iint_{-\infty < y \le x < \infty} F(dx)F(dy) = \int_{y \in \mathbb{R}} \left(\int_{x \ge y} F(dx)\right) F(dy) \\
&= \int_{y \in A} \left(\int_{x \ge y} F(dx)\right) F(dy) + \int_{y \in A^c} \left(\int_{x \ge y} F(dx)\right) F(dy) \\
&= \sum_{y \in A} \left((1 - F(y)) + F(\{y\})\right) F(\{y\}) + \int_{y \in A^c} (1 - F(y))F(dy) \\
&= \sum_{y \in A} (1 - F(y))F(\{y\}) + \int_{y \in A^c} (1 - F(y))F(dy) + \sum_{y \in A} F^2(\{y\}) \\
&= \int_{y \in \mathbb{R}} (1 - F(y))F(dy) + \sum_{y \in A} F^2(\{y\}) \\
&= 1 - E(F(X)) + \sum_{y \in A} F^2(\{y\}).
\end{aligned}$$

Summarizing we see that

$$E(F(X)) = 1 - E(F(X)) + \sum_{y \in A} F^2(\{y\})$$

and therefore

$$2E(F(X)) = 1 + \sum_{y \in A} F^2(\{y\})$$

or

$$E(F(X)) = \frac{1}{2}\left(1 + \sum_{y \in A} F^2(\{y\})\right).$$

5.10.6. (e) Note that

$$\begin{aligned}
\left|\int_{A_n} X dP - \int_A X dP\right| &= |E(X1_{A_n}) - E(X1_A)| = |E(X(1_{A_n} - 1_A))| \\
&\le E(|X||1_{A_n} - 1_A|) = E(X1_{A_n \triangle A}) \to 0
\end{aligned}$$

by the result in (b).

(a) Since

$$\left|\int_A X dP\right| \le \int_A |X| dP,$$

without loss of generality we may (and do) suppose $X \geq 0$. Now $X 1_{X>n} \to 0$ as $n \to \infty$ and

$$0 \leq X 1_{[x>n]} \leq X \in L_1$$

so by the dominated convergence theorem

$$\int_{X>n} X dP = E X 1_{X>n} \to E0 = 0.$$

(b) If $P(A_n) \to 0$, then assuming $X \geq 0$ we have for any large M

$$\begin{aligned}\int_{A_n} X dP = & \int_{A_n[X \leq M]} X dP + \int_{A_n[X>M]} X dP \\ \leq & M P(A_n[X \leq M]) + \int_{[X>M]} X dP \\ \leq & M P(A_n) + \int_{X>M} X dP.\end{aligned}$$

Therefore

$$\limsup_{n\to\infty} \int_{A_n} X dP \leq 0 + \int_{X>M} X dP.$$

Let $M \to \infty$ and by (a) we have $\int_{X>M} X dP \to 0$.

(c) If $P(A \cap [|X| > 0]) = 0$, then with $A_n = A \cap [|X| > 0]$ we have

$$\int_A |X| dP = \int_{A_n} |X| dP$$

and since $P(A_n) = 0$ we apply (b) and get

$$0 = \lim_{n\to\infty} \int_{A_n} |X| dP = \int_A |X| dP.$$

Conversely, it suffices to show that $E(|X|) = 0$ implies $P[|X| = 0] = 1$ (since then one can replace X by $X 1_A$). Recall

$$|X| \geq \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{|X| \in [\frac{k-1}{2^n}, \frac{k}{2^n})]} + n 1_{[|X|>n]}$$

so taking expectations gives

$$0 = E(|X|) \geq \sum_{k=1}^{n2^n} \frac{k-1}{2^n} P[|X| \in [\frac{k-1}{2^n}, \frac{k}{2^n})] + n P[|X| > n].$$

Therefore for $2 \leq k \leq n2^n$

$$P[[|X| \in [\frac{k-1}{2^n}, \frac{k}{2^n})] = 0 = P[|X| > n]$$

and we get by summing over k and adding the last term that $P[|X| > 2^{-n}] = 0$. Let $n \to \infty$ to get $P[|X| > 0] = 0$.

5.10.7. We have $|X_n - X| \le 2K$ and the constant function $2K$ is in L_1. Therefore by dominated convergence, $|X_n - X| \to 0$ implies that $E|X_n - X| \to 0$.

5.10.9. We have

$$\begin{aligned}\int_{-\infty}^{\infty} F(x+c) - F(x-c)dx &= \int_{\mathbb{R}} \left[\int_{u \in (x-c, x+c]} F(du) \right] dx \\ &= \int_{x-c<u<x+c} d(F \times \lambda) \\ &= \int_u \left(\int_{u-c<x<u+c} \lambda(dx) \right) F(du) \\ &= \int 2cF(du) = 2c.\end{aligned}$$

5.10.10. Define

$$X_n^* = \sum_{k-1}^{\infty} \frac{k}{2^n} 1_{[\frac{k-1}{2^n} \le X < \frac{k}{2^n}]} + \infty \cdot 1_{[X=\infty]}.$$

Note $X_n^* \ge X$, X_n^* is non-increasing and for ω such that $X(\omega) < \infty$ we have

$$\sup_{\omega \in [X<\infty]} (X_n^*(w) - X(w)) \le 2^{-n}.$$

If $E(X) = \infty$, then $\infty = E(X) \le E(X_n^*)$. If $E(X) < \infty$, then $P[X = 0] = 0$ and then

$$E(X_n^*) = E((X_n^* - X) + X) = E(X_n^* - X) + E(X) \to E(X),$$

since $0 \le E(X_n^* - X) \le 2^{-n} \to 0$.

5.10.11. Since $X \in \sigma(X)$ and $1_{[Y \in B]} \in \sigma(Y)$ we have

$$\int_{[Y \in B]} XdP = EX1_{[Y \in B]} = E(X)E(1_{[Y \in B]})$$

(from, for example, Problem 4.6.15)

$$= E(X)P[Y \in B].$$

5.10.12. (a) We first verify that $\mathcal{B} \times \mathcal{B}$ is generated by vertical and horizontal lines:

$$\mathcal{B} \times \mathcal{B} = \sigma(\{\{x\} \times X, X \times \{x\}, \forall x \in X\}) := \mathcal{X}. \qquad (*)$$

To see this note that $\{x\} \times X$ is a rectangle so

$$\{x\} \times X \in \mathcal{B} \times \mathcal{B}, \quad X \times \{x\} \in \mathcal{B} \times \mathcal{B},$$

and hence

$$\mathcal{B} \times \mathcal{B} = \sigma(\text{RECTS}) \supset \sigma(\{x\} \times X, X \times \{x\}, \forall x \in X) = \mathcal{X}.$$

To get a reverse containment let $\Lambda \in$ RECT. Then $\Lambda = A \times B$ where either

(1) A is countable and B is countable.

(2) A is countable and B^c is countable.

(3) A^c is countable and B is countable.

(4) A^c is countable and B^c is countable.

For

(1) $A \times B = \bigcup_{\substack{x \in A \\ y \in B}} \{(x,y)\} \in \mathcal{X}$.

(2) $A \times B = \bigcup_{x \in A} \{x\} \times B = \bigcup_{x \in A} \left[\{x\} \times X \setminus \bigcup_{y \in B^c} \{(x,y)\} \right] \in \mathcal{X}$.

(3) We use an argument similar to the one used in (2).

(4) $(A \times B)^c = A^c \times X \cup B^c \times X \in \mathcal{X}$.

So RECT $\subset \mathcal{X}$ and $\sigma($ RECT $) \subset \mathcal{X}$.

Now combine $(*)$ and Exercise 2.6.12 to get the following statement: If $E \in \mathcal{B} \times \mathcal{B}$, then there exists a countable set $S \subset X$ such that

$$E \in \sigma((\{s\} \times X, X \times \{s\}, s \in S) =: \mathcal{F}.$$

Let $\mathcal{P} = \{\{s\}, s \in S;\ S^c\}$ so that $\mathcal{P}$ is a partition of Ω. Then

$$\mathcal{P} \times \mathcal{P} := \{\Lambda_1 \times \Lambda_2 : \Lambda_i \in \mathcal{P}, i = 1,2\}$$

is a partition of $X \times X$ and

$$\mathcal{F} = \sigma(\mathcal{P} \times \mathcal{P}) = \Big\{ \bigcup_{\substack{j \in I, \\ k \in I'}} \Lambda_j \times \Lambda_k : I \subset \{1,2,\ldots\}, I' \subset \{1,2,\ldots\} \Big\},$$

where the last equality follows because $\mathcal{P} \times \mathcal{P}$ partitions $X \times X$.

So if $E \in \mathcal{F}$, then

$$E = \bigcup_{j,k} \Lambda_j \times \Lambda_k,$$

where j, k range over a subset of integers. But $\Lambda_j \times \Lambda_k = \{(s_i, s_j)\}$ or $\{s_i\} \times S^c$ or $S^c \times \{s_j\}$ or $S^c \times S^c$. If $E =$ DIAG then $E \in \mathcal{F}$ is impossible and we have a contradiction.

5.10.14. We proceed by means of a series of steps to show

$$Eg(X) = Eh(X, Y). \tag{\#}$$

STEP 1. If $h(x, y) = h_1(x)h_2(y)$ then $g(x) = h_1(x)E(h_2(Y))$ and

$$E(g(X)) = E(h_1(X))E(h_2(Y)).$$

Thus (#) holds. It also follows that (#) holds for $h(x, y) = 1_A(x)1_B(y) = 1_{A\times B}(x, y)$ where $A, B \in \mathcal{B}(\mathbb{R})$.

STEP 2. Let

$$\mathcal{G} := \{\Lambda \in \mathcal{B}(\mathbb{R}^2) : (\#) \text{ holds for } h = 1_\Lambda\}.$$

Note the following properties of $\mathcal{G}$:

1. $\mathbb{R}^2 \in \mathcal{G}$.
2. If $\Lambda \in \mathcal{G}$, then $\Lambda^c \in \mathcal{G}$ since
$$1 - Eg(X) = 1 - E1_\Lambda(X, Y) = E(1 - 1_\Lambda(X, Y)).$$
3. $\mathcal{G}$ is closed under countable, disjoint unions.

We conclude $\mathcal{G}$ is a λ-system. Also $\mathcal{G}$ contains $\mathcal{S} = \{A \times B : A, B \in \mathcal{B}(\mathbb{R})\}$, the π-system of measurable rectangles. Since Step 1 implies

$$\mathcal{G} \supset \mathcal{S},$$

we get from Dynkin's theorem that

$$\mathcal{G} \supset \sigma(\mathcal{S}) = \mathcal{B}(\mathbb{R}^2).$$

We conclude that (#) holds for 1_Λ whenever $\Lambda \in \mathcal{B}(\mathbb{R}^2)$.

STEP 3. Thus (#) holds for all positive simple functions and for all positive measurable functions.

5.10.15. (a) We have

$$\begin{aligned} nE\left(\frac{1}{X}1_{[X>n]}\right) &= E\left(\frac{n}{X}1_{[1>\frac{n}{X}]}\right) \le P[1 > \frac{n}{X}] \\ &= P[X > n] \to 0, \quad n \to \infty. \end{aligned}$$

(b) Now we have for any $\eta > 0$,

$$\begin{aligned} n^{-1}E\left(\frac{1}{X}1_{[X>\frac{1}{n}]}\right) =& E\left(\frac{1}{nX}1_{[1>\frac{1}{nX},X>0]}\right) \\ =& E\left(\frac{1}{nX}1_{[1>\frac{1}{nX}>\eta,X>0]}\right) + E\left(\frac{1}{nX}1_{[0<\frac{1}{nX}]}\leq \eta]\right) \\ =& A+B. \end{aligned}$$

For A we have the bound

$$A \leq P[1 > \frac{1}{nX} > \eta, X > 0] = P[1 \leq nX \leq \eta^{-1}, X > 0] \to 0$$

as $n \to \infty$. It should be clear that $B \leq \eta$ and since η is arbitrary, we are done.

5.10.16. (b) If X_1 and X_2 are independent, then $f_1(X_1)$ and $f_2(X_2)$ are independent and

$$Ef_1(X_1)f_2(X_2) = Ef_1(X_1)Ef_2(X_2).$$

Conversely, suppose $Ef(X_1)g(X_2) = Ef(X_1)Eg(X_2)$ for all bounded continuous f, g. Let $f_i = 1_{(a_i,b_i],i=1,2}$. Choose bounded continuous $f_n^{(i)}$ as in (a) and then it follows that

$$Ef_n^{(1)}(X_1)f_n^{(2)}(X_2) = Ef_n^{(1)}(X_1)Ef_n^{(2)}(X_2).$$

Let $n \to \infty$ and use the dominated convergence theorem to get

$$Ef_1(X_1)f_2(X_2) = Ef_1(X_1)Ef_2(X_2)$$

which is equivalent to

$$P[X_1 \in (a_1, b_1], X_2 \in [a_2, b_2]] = P[X_1 \in [a_1, b_1]]\, P[X_2 \in [a_2, b_2]]$$

This suffices for independence.

(c) By two applications of dominated convergence, we get for f_1, f_2 bounded and continuous

$$\begin{aligned} E(f_1(\xi_\infty)f_2(\eta_\infty)) =& \lim_{n\to\infty} E(f_1(\xi_n)f_2(\eta_n)) = \lim_{n\to\infty} E(f_1(\xi_n))E(f_2(\eta_n)) \\ =& E(f_1(\xi_\infty))E(f_2(\eta_\infty)). \end{aligned}$$

5.10.18. The Riemann integral over A would give us the area of A. Write

$$\begin{aligned} \iint_A d(\lambda \times \lambda) =& \iint 1_A d(\lambda \times \lambda) \\ =& \int_{[0,1]} \left[\int_{[0,1]} 1_{\{(x,y)\in A\}}(y)\lambda(dy)\right] \lambda(dx) \\ =& \int_{[0,1]} \left[\int_0^1 1_{A_x}(y)dy\right] \lambda(dx) \end{aligned}$$

(where the inner integral is interpreted as a Riemann integral)

$$= \int_{[0,1]} l(x)\lambda(dx)$$

where $l(x)$ is the length of a vertical line segment which passes through A at x. Note $l(x)$ is bounded, continuous and hence Riemann integrable so the above equals $\int_0^1 l(x)dx$, the area of A.

5.10.20. (a) Use the transformation theorem.

(b) If $\phi(\lambda) = 0$ then $E(e^{\lambda X}) = 0$, which implies $\exp\{\lambda X\} = 0$ almost surely. This means that either $X = -\infty$ almost surely if $\lambda > 0$ or $X = +\infty$ almost surely if $\lambda < 0$. So assuming X is $\mathbb{R}$-valued as is usual, we get $\phi(\lambda) > 0$.

Now suppose $\lambda \in \Lambda^0$. Pick ε such that $[\lambda-\varepsilon, \lambda+\varepsilon] \subset \Lambda$. Suppose $\lambda_n \to \lambda$. For all large n

$$e^{\lambda_n X} \leq e^{(\lambda-\varepsilon)X} + e^{(\lambda+\varepsilon)X} \in L_1(F)$$

and since $\lambda_n \to \lambda$, $e^{\lambda_n X} \to e^{\lambda X}$ and by dominated convergence we get $Ee^{\lambda_n X} \to Ee^{\lambda X}$.

(c) Let

$$f(x) = c(1+x)^{\alpha} e^{-x}, \quad x > 0.$$

Then $\lambda_\infty = 1$. If $\alpha < -1$, then $\int_0^\infty (1+x)^\alpha dx < \infty$ so $\lambda_\infty \in \Lambda$. If $\alpha \geq -1$, then $\int_0^\infty (1+x)^\alpha dx = \infty$, and thus $\lambda_\infty \notin \Lambda$.

(d) The density is

$$f_\lambda(x) = \frac{e^{\lambda x} f(x)}{\phi(\lambda)}.$$

(e) We have

$$\begin{aligned} F_\lambda(I) &= \int_I \frac{e^{\lambda x} F(dx)}{\phi(\lambda)} \\ &\leq \sup_{x \in I} \frac{e^{\lambda x}}{\phi(\lambda)} F(I) = 0. \end{aligned}$$

5.10.22. We use Fubini's theorem to interchange the order of integration:

$$\begin{aligned} \int_{[0,\infty]} P[X > t]dt &= \int_{[0,\infty]} \left[\int_\Omega 1_{(t,\infty)}(X(\omega))dP \right] dt \\ &= \int_{[0,\infty]\times\Omega} 1_{(t,\infty)}(X(\omega)) P \times \lambda(d\omega, dt) \\ &= \int_\Omega \left[\int_{[0,\infty)} 1_{(t,\infty)}(X(\omega))dt \right] dP \\ &= \int_\Omega X(\omega) dP(\omega) = E(X). \end{aligned}$$

5.10.25. (a) If $\gamma_n \to \gamma$, then since g is continuous, we have $g(X - \gamma_n) \to g(X - \gamma)$. Since g is bounded we get by dominated convergence that

$$\phi(\gamma_n) = Eg(X - \gamma_n) \to Eg(X - \gamma) = \phi(\gamma).$$

(b) We have

$$\lim_{\gamma\to\infty} g(X - \gamma) = g(-\infty) = -1.$$

Apply dominated convergence to get $\phi(\gamma) \to -1$ as $\gamma \to \infty$.

(c) If $\gamma_1 < \gamma_2$, then $X - \gamma_1 > X - \gamma_2$ and since g is increasing we get $g(X - \gamma_1) > g(X - \gamma_2)$ and $\phi(\gamma_1) \geq \phi(\gamma_2)$. In fact ϕ is strictly monotone: If $\gamma_1 < \gamma_2$ and $\phi(\gamma_1) = \phi(\gamma_2)$ then $0 = E(g(X - \gamma_1) - g(X - \gamma_2))$ and since the integrand is non-negative, we get $g(X - \gamma_1) - g(X - \gamma_2) = 0$ almost surely. Since g is strictly monotone we get a contradiction.

This shows (d) since $\phi(\gamma) = 0$ must have a unique root.

(e) To show $\gamma(X + c) = \gamma(X) + c$, note $\gamma(X + c)$ is the unique root of

$$E(g(X + c - \gamma) = 0,$$

which means it is the root of

$$E\Big(g(X + c - \gamma(X + c))\Big) = 0.$$

Also,

$$E\Big(g(X + c) - \gamma(X) - c)\Big) = E\Big(g(X - \gamma(X))\Big) = 0,$$

so by uniqueness, $\gamma(X + c) = \gamma(X) + c$.

(f) If $g(-x) = -g(x)$, then $\gamma(-X)$ is the root of

$$0 = Eg(-X - \gamma(-X)) = -Eg(X + \gamma(-X))$$

and since also $Eg(X - \gamma(X)) = 0$, we get by uniqueness that

$$\gamma(X) = -\gamma(-X).$$

5.10.30. We have $Y_n - X_n \geq 0$ so using Fatou's lemma

$$\begin{aligned} E(Y) - E(X) =& E(Y - X) = E\big(\liminf_{n\to\infty}(Y_n - X_n)\big) \\ \leq& \liminf_{n\to\infty} E(Y_n - X_n) = \liminf_{n\to\infty} E(Y_n) - E(X_n) \\ =& \liminf_{n\to\infty} E(Y) - E(X_n). \end{aligned}$$

Therefore

$$\limsup_{n\to\infty} E(X_n) \leq E(X)$$

and again applying Fatou's lemma we get

$$E(X) = E(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} E(X_n) \leq \limsup_{n\to\infty} E(X_n) \leq E(X).$$

5.10.31. (c) Suppose $I = [a, b]$ and $P[X \in I] \geq 1/2$. For any $\epsilon > 0$, $P[X \leq a - \epsilon] \leq 1/2$ and therefore $a - \epsilon$ cannot be a median. Similarly $b + \epsilon$ cannot be a median.

(d) Observe that

$$\begin{aligned} P[X \in & [E(X) - \sqrt{2\mathrm{Var}(X)}, E(X) + \sqrt{2\mathrm{Var}(X)}]] \\ &= P[|X - E(X)| \leq \sqrt{2\mathrm{Var}(X)}] \end{aligned}$$

and therefore by Chebychev's inequality

$$P[|X - E(X)| > \sqrt{2\mathrm{Var}(X)}] \leq \frac{\mathrm{Var(X)}}{(\sqrt{2\mathrm{Var}(X)})^2} = \frac{1}{2}.$$

Thus if

$$I = [E(X) - \sqrt{2\mathrm{Var}(X)}, E(X) + \sqrt{2\mathrm{Var}(X)}],$$

we have $P[X \in I] \geq 1/2$ and by (c), a median is in I.

5.10.36. Suppose that $X_n \in L_1$, $X_n \uparrow X$, and $\vee_n E(X_n) < \infty$.. We first show that $X \in L_1$. Since $X_n \uparrow X$, we have $X_n^- \geq X_{n+1}^- \geq X^-$, so $X_1 \in L_1$ implies $E(X^-) \leq E(X_1^-) < \infty$.

Also

$$E(X^+) = E(\liminf_{n\to\infty} X_n^+) \leq \liminf_{n\to\infty} E(X_n^+)$$

(by Fatou's lemma)

$$\begin{aligned} &= \liminf_{n\to\infty} [E(X_n^+) - E(X_n^-) + E(X_n^-)] \\ &= \liminf_{n\to\infty} [E(X_n) + E(X_n^-)] \\ &\leq \bigvee_n E(X_n) + E(X_1^-) < \infty. \end{aligned}$$

Thus $X \in L_1$ and $0 \leq X - X_n \leq X \in L_1$ and $X - X_n \to 0$ imply, by dominated convergence, that

$$E(X - X_n) = E(X) - E(X_n) \to 0.$$

Chapter 6 Solutions

6.7.1. (a) If $\{X_n(\omega)\}$ is monotone, $\lim_{n\to\infty} X_n(\omega)$ exists. Call the limit $X(\omega)$. If $X_n \xrightarrow{P} X$ there exists a sequence $\{n_j\}$ such that $X_{n_j}(\omega) \to X(\omega)$ for almost all ω. Thus $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for almost all ω.

(b) From the definition of convergence, $X_n(\omega) \to X(\omega)$ iff

$$\xi_n(\omega) = \sup_{k\geq n} |X_k(\omega) - X(\omega)| \to 0.$$

However since $\{\xi_n(\omega)\}$ is monotone, we may use (a).

(c) Since $\{Y_n\}$ is non-increasing, we need only show convergence in probability. Let the n points be $\{\exp\{2\pi i\theta_j\}, j = 1,\ldots,n\}$ where $\{\theta_j, 1 \leq j \leq n\}$ are iid U(O,1). Then

$$\begin{aligned}
[Y_n > \varepsilon] \subset & [\text{ there is an arc of length at most } 1-\varepsilon \\
& \quad \text{such that } n \text{ points are in it }] \\
\subset & [\text{ within a } (1-\varepsilon)\text{-neighborhood of some point,} \\
& \quad \text{there are } n-1 \text{ points }] \\
\subset & \bigcup_{j=1}^{n} [\text{ within a } (1-\varepsilon)\text{-neighborhood of } e^{2\pi i\theta_j}, \\
& \quad \text{there are } (n-1) \text{ points }].
\end{aligned}$$

Therefore,

$$\begin{aligned}
P[Y_n > \varepsilon] \leq & nP[\text{ within a } (1-\varepsilon)\text{-neighborhood of } e^{2\pi i\theta_1}, \\
& \quad \text{there are } (n-1) \text{ points }] \\
= & n(1-\varepsilon)^{n-1} \to 0.
\end{aligned}$$

(d) We have $\{M_n\}$ non-decreasing, so it suffices to show convergence in probability. For $x < x_0$,

$$P[M_n \leq x] = F^n(x) \to 0,$$

since $F(x) < 1$.

6.7.2. We have by the weak law of large numbers (which only requires existence of the first moment as will be shown in Chapter 7)

$$\frac{1}{n}\left(\sum_{i=1}^{n}(X_i - \bar{X}^2)\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 \xrightarrow{P} E(X_1^2) - (EX_1)^2 = \sigma^2.$$

6.7.4. We have that $\{S_n\}$ is L_2 convergent iff $\{S_n\}$ is L_2 cauchy iff

$$||S_n - S_m||_2^2 = \mathrm{Var}(\sum_{i=m+1}^{n} a_i X_i) = \sigma^2 \sum_{i=m+1}^{n} a_i^2 \to 0,$$

as $m, n \to 0$. The last statement is true iff $\{\sum_{j=1}^n a_j^2\}$ is cauchy which holds iff $\{\sum_{j=1}^n a_j^2, n \geq 1\}$ is convergent.

6.7.5. Given $\{X_n\}$ iid, $X_n \in L_1$, we show that $\{S_n/n\}$ is ui as follows. Note first that

$$\sup_{n\geq 1} E(|\frac{S_n}{n}|) \leq E(|X_1|) < \infty.$$

Next note that

$$E(|X_i|1_{[|X_i|>a]}) = E(|X_1|1_{[|X_1|>a]}) \to 0,$$

and therefore

$$\sup_{n\geq 1} E(|X_n|1_{[|X_n|>a]}) \to 0.$$

Thus $\{X_i\}$ is ui. So given $\varepsilon > 0$, there exists $\delta > 0$ such that if $P(A) < \delta$, then

$$\sup_i \int_A |X_i| dP < \varepsilon.$$

Therefore, given $\varepsilon > 0$, if $P(A) < \delta$,

$$\sup_n \int_A \left|\frac{S_n}{n}\right| dP = \sup_n \frac{1}{n}\sum_{i=1}^n \int_A |X_i| dP \leq \frac{1}{n}\sum_{i=1}^n \varepsilon = \frac{n\varepsilon}{\varepsilon} = \varepsilon.$$

We conclude $\{S_n/n\}$ is ui.

6.7.6. First of all we have

$$\sup_n E(|X_n - X|) \leq \sup_n E(|X_n|) + E(|X|) < \infty,$$

since $\{X_n\}$ is ui.

Next, suppose we are given $\varepsilon > 0$. There exists $\delta > 0$ such that if $P(A) < \delta$ then

$$\int_A |X_n| dP < \frac{\varepsilon}{2}, \qquad \int_A |X| dP < \frac{\varepsilon}{2}.$$

So if $P(A) < \delta$,

$$\int_A |X_n - X| dP \leq \int_A |X_n| dP + \int_A |X| dP \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

6.7.7. We show $\{X_n\}$ is ui iff $\sup_n \sigma_n < \infty$, where of course, $\sigma_n^2 = \mathrm{Var}(X_n)$. Suppose $N \sim N(0,1)$; that is, N has a standard normal distribution. Then

$$X_n \stackrel{d}{=} \sigma_n N.$$

If $\{X_n\}$ is ui, then

$$\infty > \sup_n E(|X_n|) = \sup_n \sigma_n E(|N|),$$

which implies the condition $\sup_n \sigma_n < \infty$ is necessary for uniform integrability. If $\sup_n \sigma_n < \infty$ is assumed, then $\sup_n E(|X_n|) < \infty$ and

$$\sup_n \int_A |X_n| dP = \sup_n \sigma_n \int_A |N| dP.$$

Given $\varepsilon > 0$, choose $\delta > 0$ so small that for $P(A) < \delta$,

$$\int_A |N| dP \leq \varepsilon / \sup_n \sigma_n.$$

6.7.9. How to get equality in the Schwartz inequality: With

$$t = E(XY)/E(Y^2)$$

we must have equality on (6.14) on page 186 so $0 = E\big((X - ty)^2\big)$ yields

$$1 = P[X - ty = 0] = P[X = \frac{E(XY)}{EY^2} \cdot Y].$$

6.7.13. We have $E(|X_1|^2) < \infty$ and

$$nP[|X_1| > \varepsilon\sqrt{n}] = E\Big(n 1_{[\frac{|X_1|^2}{\varepsilon^2} \geq n]}\Big) \leq \frac{1}{\varepsilon^2} E\Big(|X_1|^2 1_{[|X_1|^2 > \varepsilon^2 n]}\Big) \to 0,$$

since $|X_1|^2 \in L_1$.

Then

$$P\left[\bigvee_1^n \frac{|X_k|}{\sqrt{n}} > \varepsilon\right] = P\Big\{\bigcup_{k=1}^n [|X_k| > \varepsilon\sqrt{n}]\Big\} \leq nP[|X_1| > \varepsilon\sqrt{n}] \to 0.$$

6.7.15. Write

$$\begin{aligned} E(|X_0 - X_n|) \leq & E\big((X_0 - X_n) 1_{[X_0 \geq X_n]}\big) + E\big((X_n - X_0) 1_{[X_n \geq X_0]}\big) \\ = & A + B. \end{aligned}$$

For A:

$$(X_0 - X_n) 1_{[X_0 \geq X_n]} \leq X_0 \in L_1$$

and

$$P[(X_0 - X_n) 1_{[X_0 \geq X_n]} > \epsilon] \leq P[|X_0 - X_n| > \epsilon] \to 0.$$

Thus, by dominated convergence,

$$A = E\big((X_0 - X_n) 1_{[X_0 \geq X_n]}\big) \to 0.$$

For B, use a variant of Pratt's lemma (Problem 5.10.30): if

$$0 \le \xi_n \le \eta_n$$

and

$$\xi_n \xrightarrow{P} \xi_\infty, \quad \eta_n \xrightarrow{P} \eta_\infty,$$

and $E(\eta_n) \to E(\eta_\infty) < \infty$, then $E(\xi_n) \to E(\xi_\infty) < \infty$ as well. To see this, let $\{E(\xi_{n'})\}$ be a convergent subsequence. There exists a further subsequence $\{n''\}$ such that along this subsequence both

$$\xi_{n''} \xrightarrow{\text{a.s.}} \xi_\infty, \quad \eta_{n''} \xrightarrow{\text{a.s.}} \eta_\infty.$$

By Pratt, $E(\xi_{n''}) \to E(\xi_\infty)$ and hence any convergent subsequence of $\{E(\xi_n)\}$ converges to the correct limit; therefore the full sequence converges as well.

Back to B: We have

$$0 \le (X_n - X_0)1_{[X_n \ge X_0]} \le X_n$$

and $E(X_n) \to E(X_0)$ and

$$P[(X_n - X_0)1_{[X_n \ge X_0]} > \epsilon] \to 0.$$

Thus

$$E\Big((X_n - X_0)1_{[X_n \ge X_0]} > \epsilon]\Big) \to 0$$

by the Pratt lemma variant.

Thus we conclude $E(|X_n - X_0|) \to 0$ as required.

6.7.16. (a) If a sequence converges to 0, then its Cesaro averages converge to 0.

(b) If $\|X_n\|_p \to 0$, then by Minkowski (triangle) inequality

$$\left\|\frac{\sum_{i=1}^n X_i}{n}\right\|_p \le \frac{1}{n}\sum_{i=1}^n \|X_i\|_p \to 0$$

since convergence to 0 always implies Cesaro convergence to 0.

(c) Let the probability space be $[0,1]$ with Lebesgue measure. Define X_1, X_2 to be the indicators of $(0, 1/2], (1/2, 1]$ so that $X_1 + X_2 = 2$. Then define X_3, X_4, X_5 to be the indicators of the three subintervals of $(0,1]$ of length $1/3$ so that $X_3 + X_4 + X_5 = 3$. Let $X_6, \dots, X_9$ be the indicators of the 4 subintervals of length $1/4$ so that $X_6 + \dots + X_9 = 4$ and so on. Then $X_n \xrightarrow{P} 0$ since the length of the intervals on which any indicator is different from 0 shrinks. However

$$\begin{aligned}
\frac{1}{2}(X_1 + X_2) =& 1,\\
\frac{1}{5}(X_1 + \dots + X_5) =& \frac{2+3}{5} = 1,\\
\frac{1}{9}(X_1 + \dots + X_9) =& \frac{2+3+4}{9} = 1,
\end{aligned}$$

and so on. Therefore $\frac{1}{n}\sum_1^n X_i$ does not converge in probability to 0.

(d) We write

$$\frac{X_n}{n} = \frac{S_n}{n} - \Big(\frac{n-1}{n}\Big)\frac{S_{n-1}}{n-1} \xrightarrow{P} 0 - (1)\cdot 0 = 0.$$

6.7.19. We have for any $\delta > 0$

$$P[|X_n| > \delta] \le P[Y_n > \delta] \to 0,$$

since $Y_n \xrightarrow{P} 0$.

6.7.20. Suppose Y is a non-negative random variable satisfying

$$E(Y) = 1, \quad E(Y^2) = b > 0.$$

Further suppose $0 < a < 1$ and define

$$u(x) = \frac{(x-a)(a+\frac{2b}{1-a}-x)}{(b/(1-a))^2}.$$

Then $u(\cdot)$ is a quadratic function with roots at a and $a + 2b/(1-a)$ and with a positive maximum of 1 at the argument $a + b/(1-a)$. Note further that

$$u(x) \le \begin{cases} 0, & \text{if } x \le a \text{ or } x > a + 2b/(1-a), \\ 1, & \text{if } x \in \mathbb{R}. \end{cases}$$

On the one hand,

$$\begin{aligned} Eu(Y) =& Eu(Y)1_{[Y\notin[a,a+2b/(1-a)]]} + Eu(Y)1_{[Y\in[a,a+2b/(1-a)]]} \\ \le& 0\cdot P[Y \notin [a, a+2b/(1-a)]] + 1\cdot P[Y \in [a, a+2b/(1-a)]] \\ \le& P[Y \ge a]. \end{aligned}$$

On the other hand,

$$\begin{aligned} Eu(Y) =& \frac{(1-a)^2}{b^2}E\Big((Y-a)(a+\frac{2b}{1-a}-Y)\Big) \\ =& \frac{(1-a)^2}{b^2}\{a + \frac{2b}{1-a} - b - a(a+\frac{2b}{1-a}) + a\} \\ =& \frac{(1-a)^2}{b^2}\{(a+\frac{2b}{1-a})(1-a) - b + a\} \\ =& \frac{(1-a)^2}{b^2}\{a(1-a) + 2b - b + a\} \\ =& \frac{(1-a)^2}{b^2}\{2a - a^2 + b\}. \end{aligned}$$

For $0 < a < 1$, $2a - a^2 \geq 0$ and so the above is bounded below by

$$\geq \frac{(1-a)^2}{b^2} \cdot b = \frac{(1-a)^2}{b}.$$

We conclude that

$$P[Y \geq a] \geq \frac{(1-a)^2}{b}.$$

Now for X satisfying $E(X^2) = 1$ and $E(|X|) \geq a > 0$, set $Y = |X|/E(|X|)$. Then

$$E(Y) = 1, \quad E(Y^2) = \frac{E(X^2)}{E^2(|X|)} = \frac{1}{E^2(|X|)} =: b.$$

For $0 < \lambda < 1$,

$$P[|X| \geq \lambda a] = P\Big[Y \geq \frac{\lambda a}{E(|X|)}\Big] = P[Y \geq a'],$$

for $a' = \lambda a/E(|X|) \in (0,1)$, since $0 < a/E(|X|) < 1$. Thus

$$\begin{aligned} P[|X| \geq \lambda a] =& P[Y \geq a'] \geq \frac{(1-a')^2}{E(Y^2)} \\ =& \Big(1 - \frac{\lambda a}{E(|X|)}\Big)^2 \cdot E^2(|X|) \\ =& (E(|X|) - \lambda a)^2 = a\Big(\frac{E(|X|)}{a} - \lambda\Big)^2 \geq a(1-\lambda)^2. \end{aligned}$$

6.7.23. A sequence of random variables $\{X_n\}$ converges in probability to ∞ if for any M, we have

$$P[X_n \geq M] \to 1, \quad n \to \infty.$$

For any integer M,

$$P[T(s) \geq M] = \sum_{k \geq M} (1-s)s^k = s^M \to 1,$$

as $s \to 1$. So $T(s) \xrightarrow{P} \infty$.

Note

$$(1-s)U(s) = \sum_{n=0}^{\infty} (1-s)s^n u_n = E(u_{T(s)}).$$

Now $T(s) \xrightarrow{P} \infty$ as $s \to 1$, implies $u_{T(s)} \xrightarrow{P} u$. To see this, observe that given any $\delta > 0$, there exists n_0 such that

$$|u_n - u| \leq \delta.$$

Therefore

$$P[|u_{T(s)} - u| > \delta] \leq P[T(s) \leq n_0] \to 0,$$

as $s \to 1$. The result now follows by applying dominated convergence to convergence in probability.

6.7.24. (a) We have for any $\delta > 0$

$$\begin{aligned} P[|X_n - X|^2 > \delta^2] \leq & P[|X_n - X|^2 + |Y_n - Y|^2 > \delta^2] \\ = & P[d((X_n, Y_n), (X, Y)) > \delta] \to 0. \end{aligned}$$

(b) By part (a), it suffices to assume the range of f is $\mathbb{R}$.

Given any subsequence $\{n(k)\}$, there exists a further subsequence $\{n(k')\} \subset \{n(k)\}$ such that

$$X_{n(k')} \to X, \quad Y_{n(k')} \to Y$$

almost surely and by continuity of f,

$$f(X_{n(k')}, Y_{n(k')}) \to f(X, Y)$$

almost surely. By the subsequence characterization of convergence in probability we have

$$f(X_n, Y_n) \xrightarrow{P} f(X, Y).$$

(c) Define the continuous function $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ by

$$f(x, y) = (x + y, xy).$$

Apply (b).

6.7.25. (d) Suppose it is possible to metrize almost sure convergence with the metric $d(\cdot, \cdot)$. Let $\{X_n, n \geq 1\}$ be a sequence of random variables such that $X_n \xrightarrow{P} X$ but that $\{X_n\}$ does not converge almost surely. For instance, Example 6.2.1 provides such a sequence. Since almost sure convergence fails, there exist a subsequence $\{n_k\}$ and a $\delta > 0$ such that

$$d(X_{n_k}, X) > \delta.$$

Since $X_n \xrightarrow{P} X$, given the subsequence $\{n_k\}$, there exists a further subsequence $\{n_{k(j)}\} \subset \{n_k\}$ such that almost sure convergence holds along the subsubsequence: $X_{n_{k(j)}} \xrightarrow{\text{a.s.}} X$ and therefore

$$d(X_{n_{k(j)}}, X) \to 0.$$

However, this contradicts the previous display so metrizing almost sure convergence is impossible.

6.7.26. (a) If

$$X_n = \frac{n}{\log n} 1_{(0,1/n)},$$

then

$$E(X_n) = \frac{n}{\log n} \cdot \frac{1}{n} = \frac{1}{\log n} \to 0.$$

Also we have

$$EX_n(s)1_{\{u:X_n(u)>a\}}(s) = \int_0^{1/n} \frac{n}{\log n} 1_{\{u:n/\log n>a \text{ and } u<1/n\}}(s)ds$$
$$= \begin{cases} \frac{1}{\log n}, & \text{if } \frac{n}{\log n} > a, \\ 0, & \text{if } \frac{n}{\log n} < a. \end{cases}$$

Therefore if we set $U(x) = x/\log x$ we have

$$\sup_n \int_{[|X_n|>a]} |X_n| dP = \bigvee_{n/\log n > a} \frac{1}{\log n} \le \bigvee_{n \ge U^{\leftarrow}(a)} \frac{1}{\log n} \to 0$$

as $a \to \infty$.

Finally, to see that there is no dominating variable, we suppose there is one and get a contradiction. Suppose

$$X_n \le Y \in L_1(0,1).$$

This means on $(0, 1/n)$, we have $Y \ge n/\log n$. Thus

$$E(Y) = \sum_{n=1}^{\infty} E\Big(Y(s)1_{[\frac{1}{n+1},\frac{1}{n}]}(s)\Big)$$
$$\ge \sum_{n=1}^{\infty} \frac{n}{\log n}\left(\frac{1}{n} - \frac{1}{n+1}\right)$$
$$= \sum_{n=1}^{\infty} \frac{1}{(n+1)\log n} = \infty,$$

by comparison with the integral $\int_1^\infty y^{-1}dy$.

(b) Suppose

$$X_n = n1_{(0,1/n)} - n1_{[1/n,2/n)}.$$

Then

$$E(X_n) = n \cdot \frac{1}{n} - n \cdot \frac{1}{n} = 0$$

and for any $\epsilon > 0$

$$P[|X_n| > \epsilon] \le P(0, 2/n) = 2/n \to 0.$$

Note that

$$|X_n(s)| = \begin{cases} n, & \text{if } 0 < s < \frac{2}{n} \\ 0, & \text{if } \frac{2}{n} < s < 1. \end{cases}$$

Therefore

$$E\left(|X_n|1_{[|X_n|>a]}\right) = \int_0^{2/n} n1_{\{u:n>a,u<2/n\}}(s)ds\Big) = \begin{cases} n \cdot \frac{2}{n} = 2, & \text{if } n > a, \\ 0, & \text{if } n < a, \end{cases}$$

and thus, finally,

$$\sup_n E\left(|X_n|1_{[|X_n|>a]}\right) = 2,$$

so the sequence $\{X_n\}$ is not ui.

6.7.31. (a) Since $X_n - c_n \xrightarrow{P} 0$, we have $P[|X_n - c_n| \le \epsilon] \to 1$. Thus for all large n,

$$P\big[X_n \in [c_n - \epsilon, c_n + \epsilon]\big] \ge 1/2.$$

From Exercise 5.5.31, we get that

$$m(X_n) \in [c_n - \epsilon, c_n + \epsilon]$$

for all large n. Since ϵ is arbitrary, we have $m(X_n) - c_n \to 0$ as $n \to \infty$.

(b) Because X has a unique median m, for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$P[X \le m - \epsilon] < \frac{1}{2} - \delta, \quad P[X \ge m + \epsilon] < \frac{1}{2} - \delta.$$

Therefore,

$$\begin{aligned} P[X_n \le m - \epsilon] =& P[X_n \le m - \epsilon, |X - X_n| \le \epsilon/2] \\ &+ P[X_n \le m - \epsilon, |X - X_n| > \epsilon/2] \\ \le& P[X \le m - \epsilon/2] + P[|X - X_n| \le \epsilon/2] \\ \le& \frac{1}{2} - \delta + o(1). \end{aligned}$$

Therefore, for all large n, we have $m(X_n) \ge m - \epsilon/2$. In a similar way, we show $m(X_n) \le m + \epsilon/2$.

6.7.33. Assuming that $X_{ni} \ge 0$, we have

$$\begin{aligned} P[\frac{S_n}{n} > \epsilon] =& P[S_n > n\epsilon] \le P\{\bigcup_{j=1}^{n} [X_{nj} > \epsilon]\} \\ =& P[\bigvee_{j=1}^{n} X_{nj} > \epsilon] \to 0 \end{aligned}$$

as $n \to \infty$.

CHAPTER 7 SOLUTIONS

7.7.6. We have $E(X_k) = \gamma_k$ so that $\sum_k \gamma_k < \infty$ implies $\sum_k E(X_k) < \infty$ and therefore $\sum_k X_k < \infty$ a.s.

Conversely, if $\sum_k X_k < \infty$ almost surely, then

$$0 < E\Big(e^{-\sum_k X_k}\Big) = E\Big(\prod_1^\infty e^{-X_k}\Big) = \prod_1^\infty E(e^{-X_k}) = \prod_{k=1}^\infty \left(\frac{1}{2}\right)^{\gamma_k}$$

and $\prod_k \left(\frac{1}{2}\right)^{\gamma_k} > 0$ iff $\sum_k \gamma_k(\log 2) < \infty$ iff $\sum_k \gamma_k < \infty$.

7.7.7. (a) If $1_{\cup_{k=1}^n E_k}(\omega) = 1$, then $\omega \in \cup_{k=1}^n E_k$ and the left side equals the right side since both are 1. If $1_{\cup_{k=1}^n E_k}(\omega) = 0$, then both sides are zero.

Using (a) and the Schwartz inequality, we have

$$\begin{aligned}\left(E\sum_1^n 1_{E_k}\right)^2 &= \left(E\Big(1_{\cup_{k=1}^n E_k} \cdot \sum_{k=1}^n 1_{E_k}\Big)\right)^2 \le E\Big(1^2_{\cup_{k=1}^n E_k}\Big)E(\sum_{k=1}^n 1_{E_k})^2 \\ &= E(1_{\cup_{k=1}^n E_k})E(\sum_{k=1}^n 1_{E_k})^2 = P(\bigcup_{k=1}^n E(\sum 1_{E_k})^2.\end{aligned}$$

Therefore

$$P\left(\bigcup_{k=1}^n E_k\right) \ge \frac{\left(E\sum_1^n 1_{E_k}\right)^2}{E\left(\sum_1^n 1_{E_k}\right)^2}.$$

(b) Suppose

(i) $\sum_{n=1}^\infty PE_n = \infty$,

(ii) $P(E_mE_n) \le cP(E_m)PE(n-m)$.

Then

$$
\begin{aligned}
P(\limsup_{n\to\infty} E_n) &= \lim_{n\to\infty} P\Big(\bigcup_{j=n+1}^{\infty} E_j\Big) = \lim_{n\to\infty}\lim_{N\to\infty} P\Big(\bigcup_{j=n+1}^{\infty} E_j\Big) \\
&\geq \lim_{n\to\infty}\lim_{N\to\infty} \frac{\Big(E\sum_{j=n+1}^{N} 1_{E_j}\Big)^2}{E\Big(\sum_{j=n+1}^{N} 1_{E_j}\Big)^2} \\
&= \lim_{n\to\infty}\lim_{N\to\infty} \frac{\Big(\sum_{j=n+1}^{N} P(E_j)\Big)^2}{\sum_{j=n+1}^{N} P(E_j) + 2\sum_{n+1\leq j<k\leq N} P(E_j E_k)} \\
&\geq \lim_{n\to\infty}\lim_{N\to\infty} \frac{\Big(\sum_{n+1}^{N} P(E_j)\Big)^2}{\sum_{j=n+1}^{N} P(E_j) + 2c\sum_{n+1\leq j<k\leq N} P(E_j)P(E_{k-j})} \\
&\geq \lim_{n\to\infty}\lim_{N\to\infty} \frac{\Big(\sum_{j=n+1}^{N} P(E_j)\Big)^2}{\sum_{j=n+1}^{N} P(E_j) + 2c\Big(\sum_{j=n+1}^{N} P(E_j)\Big)^2} \\
&= \lim_{n\to\infty}\lim_{N\to\infty} \frac{1}{\big(\sum_{j=n+1}^{N} P(E_j)\big)^{-1} + 2c} = \frac{1}{2c} > 0.
\end{aligned}
$$

(c) Suppose $Y_n \geq 0$ are iid with common distribution G and $X_n \geq 0$ are iid with common distribution F. We have from Fubini's theorem

$$
\begin{aligned}
\sum_{n=1}^{\infty} P\Big[\frac{Y_n}{\bigvee_{i=1}^{n} X_i} > \epsilon\Big] &= \sum_n \iint_{\{(x,y):\frac{y}{x}>\epsilon\}} G(dy)F^n(dx) \\
&= \sum_n \int_{y\in[0,\infty)} \Big[\int_{\{x:y\varepsilon^{-1}>x\}} F^n(dx)\Big] G(dy) \\
&= \sum_n \int_0^{\infty} F^n\Big(\frac{y}{\varepsilon}\Big) G(dy) \\
&= \int \frac{G(dy)}{1-F(y/\varepsilon)}.
\end{aligned}
$$

So if

$$
\int_0^{\infty} \frac{G(dy)}{1-F(y/\varepsilon)} < \infty, \qquad \forall\, \varepsilon > 0,
$$

then

$$
\sum_{n=1}^{\infty} P\Big[\frac{Y_n}{\bigvee_{i=1}^{n} X_i} > \varepsilon\Big] < \infty
$$

and

$$P[\frac{Y_n}{\bigvee_{i=1}^{n} X_i} > \varepsilon \text{ i.o. }] = 0,$$

which implies that

$$\frac{Y_n}{\bigvee_{i=1}^{n} X_i} \to 0$$

almost surely.

For the converse, we suppose $Y_n / \bigvee_{i=1}^n X_i \stackrel{\text{a.s.}}{\to} 0$. Set $E_n = [\frac{Y_n}{\bigvee_{i=1}^n X_i} > \varepsilon]$ so that $P(E_n \text{ i.o. }) = 0$. For $m < n$

$$\begin{aligned}
P(E_m E_n) =& P\left[\frac{Y_m}{\bigvee_{i=1}^m X_i} > \varepsilon, \frac{Y_n}{\bigvee_{i=1}^{n} X_i} > \varepsilon\right] \\
\leq& P\left[\frac{Y_m}{\bigvee_{i=1}^{m} X_i} > \varepsilon, \frac{Y_n}{\bigvee_{i=m+1}^{n} X_i} > \varepsilon\right] \\
=& P(E_m)P(E_{n-m}).
\end{aligned}$$

We conclude that $\sum_{n=1}^{\infty} P(E_n) < \infty$, and it follows that

$$\int \frac{G(dy)}{1 - F(\frac{y}{\varepsilon})} < \infty.$$

7.7.8. From problem 6.7.5, we have $\{S_n/n, n \geq 1\}$ is ui and hence the SLLN plus Theorem 6.6.1 proves L_1-convergence.

7.7.9. Define $X_j' = X_j 1_{[|X_j| \leq j]}$. Then

$$\sum_j P[\frac{c_j X_j'}{j} \neq \frac{c_j X_j}{j}] \leq \sum_j P[|X_j| > j] = \sum_j P[|X_1| > j] < \infty$$

since $E|X_1| < \infty$. Therefore $\sum_j \frac{c_j X_j}{j}$ converges almost surely iff $\sum_j \frac{c_j X_j'}{j}$ converges almost surely and for this, it suffices to check

$$\sum_j \operatorname{Var}\left(\frac{c_j X_j'}{j}\right) \leq \sum_j \frac{c_j^2}{j^2} E\left(X_j^2 1_{[|X_j| \leq j]}\right) < \infty.$$

To see that the right side expression indeed converges, note that it is bounded by

$$\bigvee_k |c_k|^2 \sum_j \frac{1}{j^2} E\left(X_j^2 1_{[|X_j| \leq j]}\right) < \infty.$$

Note, $\bigvee_k |c_k|^2 < \infty$ since $\{c_j\}$ is bounded and the sum of the expectations is finite by the argument used to prove the strong law of large numbers. We therefore conclude that

$$\sum_j \frac{c_j X_j}{j} \text{converges almost surely}$$

and by the Kronecker lemma, we get

$$\sum_{j=1}^{n} \frac{c_j X_j}{n} \to 0 \text{ a.s.}$$

as $n \to \infty$.

7.7.10. (a) Start by supposing that $\{\xi_j\}$ are independent and $E(\xi_j) = 0$, and $|\xi_j| \leq M$. Then $\frac{1}{n}\sum_{j=1}^{n} \xi_j \overset{\text{a.s.}}{\to} 0$. To see this, it suffices, by the Kronecker lemma, to show that $\sum_j \frac{\xi_j}{j}$ converges almost surely. For this, it is enough to check the Kolmogorov convergence criterion that $\sum_j \mathrm{Var}(\xi_j/j) < \infty$. However, we have

$$\sum_j \mathrm{Var}(\xi_j/j) = \sum_j \frac{1}{j^2}\mathrm{Var}(\xi_j) \leq \sum_j \frac{1}{j^2} E(\xi_j^2) \leq \sum_j \frac{M^2}{j^2} < \infty.$$

Now let $\xi_j^{(i)} = X_{i+j(m+1)}$ where $\{X_n\}$ is m-dependent. From the above discussion, we have $\sum_{j=1}^{n} \xi_j^{(i)}/n \overset{\text{a.s.}}{\to} 0$ as $n \to \infty$. However,

$$\begin{aligned}
\frac{1}{n}\sum_{\alpha=1}^{n} X_\alpha &= \frac{1}{n}\sum_{i=1}^{m+1} \sum_{j: i+j(m+1)\leq n} X_{i+j(m+1)} \\
&= \sum_{i=1}^{m+1} \frac{[\frac{n-i}{m}]}{n} \left(\frac{\sum_{j=0}^{[\frac{n-i}{m+1}]} \xi_j^{(i)}}{[\frac{n-i}{m+1}]} \right) \\
&\to \sum_{i=1}^{m+1} 1 \cdot 0 = 0.
\end{aligned}$$

(b) We have

$$N_n(u_1, \ldots, u_k) = \frac{1}{n}\sum_{m=1}^{n} 1_{[X_m = u_1, \ldots, X_{m+k-1} = u_k]}.$$

The indicators are k-dependent and

$$I_{[X_m=u_1,\ldots,X_{m+k-1}=u_k]} - \prod_{i=1}^{k} p_{u_i}$$

is bounded so from (a)

$$\frac{N_n(u_1,\ldots,u_k)}{n} - \prod_{i=1}^{k} p_{u_i} \stackrel{\text{a.s.}}{\to} 0.$$

7.7.12. Let $\{U_n, n \geq 0\}$ be iid $U(0,1)$ random variables. Then $X_0 = U_0$ and $X_{n+1} = U_{n+1}X_n$ so that $X_{n+1} = \prod_{i=0}^{n+1} U_i$. Therefore, by the strong law of large numbers

$$\frac{1}{n}\log X_n = \frac{1}{n}\sum_{i=0}^{n} \log U_i \stackrel{\text{a.s.}}{\to} -1,$$

since, for $x > 0$, $P[-\log U_1 > x] = P[U_1 \leq e^{-x}] = e^{-x}$ and $-\log U_1$ has a unit exponential distribution.

7.7.13. Use Problem 7.7.15. Then $\sum_n X_n < \infty$ almost surely iff for any $c > 0$,

(1) $\sum_n P[|X_n| > c] = \sum_n e^{-\lambda_n c} < \infty$.

(2) $\sum_n E\Big(X_n 1_{[|X_n| \leq c]}\Big) = \sum_n \left[\lambda_n^{-1}(1 - e^{-\lambda_n c}) - ce^{-\lambda_n c}\right] < \infty$.

Now (1) implies $\lambda_n \to \infty$ and thus (1) and (2) hold iff

$$(2') \ \sum_n \lambda_n^{-1}(1 - e^{-\lambda_n c}) < \infty$$

and (1) hold. However the series in (2') is the same as $\sum_n \lambda_n^{-1} - \sum_n \lambda_n^{-1} e^{-\lambda_n c}$ and thus, we have that (1) and (2) hold iff (1) holds.

7.7.14. If $\sum_n \sigma_n^2 < \infty$, then by the Kolmogorov criterion, $\sum_n (X_n - \mu_n)$ converges almost surely and since $\sum_n \mu_n$ converges, we get $\sum_n X_n$ converges almost surely.

Conversely, suppose $\sum_n X_n$ converges almost surely. Let $\{X_j'\}$ be iid copies of $\{X_n\}$ and then $\sum_{j=1}^{n}(X_j - X_j')$ converges almost surely. Note

$$\operatorname{Var}\Big(\sum_{j=1}^{n}(X_j - X_j')\Big) = 2\sum_{j=1}^{n}\sigma_j^2 =: s_n^2$$

and $N_n := \sum_{j=1}^{n}(X_j - X_j')$ is a normal random variable with mean 0 and variance s_n^2. So we assume

$$N_n \Rightarrow X_\infty$$

since $\{N_n\}$ is almost surely convergent, where X_∞ is some proper random variable. For the purposes of getting a contradiction, suppose $s_n \to \infty$. Let $N(0,1)$ be a standard normal rv with mean 0 and variance 1 and then for any $x \in \mathbb{R}$ as $n \to \infty$

$$\begin{aligned} P[N_n \le x] =& P[s_n N(0,1) \le x] = P[N(0,1) \le \frac{x}{s_n}] \\ \to& P[N(0,1) \le 0] = \frac{1}{2} = P[X_\infty \le x] \end{aligned}$$

which means X_∞ cannot be proper.

Therefore, $\sum_j \sigma_j^2 < \infty$ and $\sum_j (X_j - \mu_j)$ is convergent by the Kolmogorov criterion. Since $\sum_j X_j$ is convergent, we get $\sum_j \mu_j$ is convergent.

7.7.15. We suppose that $V_n > 0$ and that

$$\sum_n P[V_n > c] < \infty, \quad \sum_n E\Big(V_n 1_{[V_n \le c]}\Big) < \infty.$$

Then it follows that

$$\sum_n \mathrm{Var}(V_n 1_{[V_n \le c]}) \le \sum_n E(V_n^2 1_{[V_n \le c]}) \le c \sum_n E(V_n 1_{[V_n \le c]}) < \infty.$$

7.7.16. Since $|S_n/n| \sim |\mu| \neq 0$ by the strong law of large numbers, it suffices to show

$$\frac{1}{n} M_n := \frac{1}{n} \bigvee_{i=1}^{n} |X_i| \stackrel{\text{a.s.}}{\to} 0,$$

as $n \to \infty$. Since $E(|X_1|) < \infty$, we have

$$\lim_{n\to\infty} \frac{1}{n} |X_n(\omega)| = 0,$$

for $\omega \in \Lambda$ and $P(\Lambda) = 1$. Now for each n, there exists $k(n) \le n$ (which is random) such that $M_n = X_{k(n)}$ and therefore,

$$\frac{1}{n} M_n \le \frac{|X_{k(n)}|}{k(n)}.$$

Suppose $\omega \in \Lambda$. There are two cases. In case 1, $k(n,\omega) \to \infty$, so that

$$\frac{1}{n} M_n(\omega) \le \frac{|X_{k(n,\omega)}(\omega)|}{k(n,\omega)} \to 0.$$

In case 2, for some integer $M(\omega) < \infty$, we have $k(n,\omega) \le M$ so that

$$\frac{1}{n} M_n(\omega) \le \frac{1}{n} \bigvee_{i=1}^{M(\omega)} |X_i(\omega)| \to 0.$$

The desired result follows for either case.

7.7.17. We have

$$\sum_k P[X_k = k^2] = \sum_k \frac{1}{k^2} < \infty$$

so $P[X_k = k^2 \text{ i.o.}] = 0$ and for $k \geq k_0(\omega)$, we have for almost all ω that $X_k(\omega) = -1$. Therefore, $\sum_k X_k = -\infty$.

7.7.18. We have that $\sum_j \frac{X_j}{j^{1/2}\log j}$ converges almost surely by the Kolmogorov convergence criterion, since

$$\sum_j \mathrm{Var}\left(\frac{X_j}{j^{\frac{1}{2}}\log j}\right) = \sum_j \frac{1}{j(\log j)^2} < \infty.$$

By Kronecker's lemma

$$\frac{\sum_{j=1}^{n} X_j}{n^{1/2}\log n} \xrightarrow{\text{a.s.}} 0.$$

7.7.21. We need the fact that $\sum_{i=1}^{n+1} X_i(\theta)$ has density

$$f_\theta(x) = \frac{1}{\theta}(x/\theta)^n e^{-x/\theta}/n!, \quad x > 0,$$

so that

$$\bar{X}_{n+1}(\theta) = \frac{1}{n+1}\sum_{i=1}^{n+1} X_i(\theta) \text{ has density } (n+1)f_\theta((n+1)x), \quad x > 0.$$

By the weak law of large numbers

$$\bar{X}_n(\theta) \xrightarrow{P} \theta$$

and by continuity we also get

$$u(\bar{X}_n(\theta)) \xrightarrow{P} u(\theta),$$

and since u is bounded, dominated convergence yields

$$E_\theta(u(\bar{X}_n(\theta))) = \int_0^\infty nf_\theta(nx)u(x)dx \to u(\theta).$$

We now check uniform convergence on a compact interval $I = [a,b]$. Let $I^+ = [a/2, 2b]$. Define the modulus of continuity of u for any compact interval J by

$$\omega_\delta(J) := \sup_{\substack{|x-y|\leq\delta \\ x,y\in J}} |u(x) - u(y)|$$

and since u is uniformly continuous on J, we have $\omega_\delta(J) \to 0$ as $\delta \downarrow 0$. Now decompose as follows:

$$\begin{aligned}
\sup_{\theta\in I} |E_\theta(u(\bar{X}_n(\theta))) - u(\theta)| \leq & \sup_{\theta\in I} E_\theta \left|u(\bar{X}_n(\theta))) - u(\theta)\right| \\
\leq & \sup_{\theta\in I} E_\theta \left|u(\bar{X}_n(\theta))) - u(\theta)\right| 1_{[|\bar{X}_n(\theta)-\theta|\leq\delta]} \\
& + \sup_{\theta\in I} E_\theta \left|u(\bar{X}_n(\theta))) - u(\theta)\right| 1_{[|\bar{X}_n(\theta)-\theta|\geq\delta]}
\end{aligned}$$

which for δ small has an upper bound

$$\begin{aligned}
&\leq \omega_\delta(I^+) + 2\|u\| \sup_{\theta\in I} P_\theta[|\bar{X}_n(\theta) - \theta| > \delta] \\
&= I + II.
\end{aligned}$$

Here, $\|u\| = \sup_{x\in\mathbb{R}} |u(x)| < \infty$. Now for II we have

$$\begin{aligned}
II \leq & 2\|u\| \frac{1}{\delta^2} \sup_{\theta\in I} \mathrm{Var}(\bar{X}_n(\theta)) = 2\|u\| \frac{1}{\delta^2} \sup_{\theta\in I} \frac{1}{n} \mathrm{Var}(X_1(\theta)) \\
= & 2\|u\| \frac{1}{\delta^2} \frac{1}{n} \sup_{\theta\in I} \theta^2 \to 0
\end{aligned}$$

as $n \to \infty$. Therefore

$$\limsup_{n\to\infty} \sup_{\theta\in I} |E_\theta(u(\bar{X}_n(\theta))) - u(\theta)| \leq \omega_\delta(I^+),$$

for any $\delta > 0$ and letting $\delta \downarrow 0$, yields the result.

7.7.22. (a) You can differentiate under the integral sign; this is justified by dominated convergence. For instance,

$$\begin{aligned}
-\Big(\frac{\hat{F}(\lambda+\delta) - \hat{F}(\lambda)}{\delta}\Big) &= \int_0^\infty \frac{1}{\delta}(1 - e^{-\delta x}) e^{-\lambda x} F(dx) \\
&= \int_0^\infty \Big(\int_0^x e^{-\delta y} dy\Big) e^{-\lambda x} F(dx).
\end{aligned}$$

Set

$$G_\lambda(dx) = e^{-\lambda x} F(dx)/\hat{F}(\lambda), \quad H_\delta(x) = \int_0^x e^{-\delta y} dy.$$

Then

$$H_\delta(x) \leq x \in L_1(G_\lambda)$$

and as $\delta \to 0$

$$H_\delta(x) \to x.$$

Therefore, by dominated convergence, as $\delta \to 0$

$$-\Big(\frac{\hat F(\lambda+\delta)-\hat F(\lambda)}{\delta}\Big) = \int_0^\infty H_\delta(x)G_\lambda(dx)\cdot \hat F(\lambda)$$
$$\to \int_0^\infty xG_\lambda(dx)\cdot \hat F(\lambda) = \int_0^\infty xe^{-\lambda x}F(dx).$$

(b) Fix θ. Then by the weak law of large numbers

$$\frac{1}{n}\sum_{i=1}^n \xi_i(\theta) \xrightarrow{P} \theta,$$

since the limit is the mean. The rest follows from the fact that $\sum_{i=1}^n \xi_i(\theta)$ has a Poisson distribution with parameter $n\theta$.

(c) Write

$$\sum_{j\le nx}\frac{(-1)^j}{j!}n^j\hat F^{(j)}(n) = \int_0^\infty \sum_{j\le nx} e^{-ns}\frac{(ns)^j}{j!}F(ds)$$
$$= \int_0^\infty P[\frac{1}{n}PO(ns)\le x]F(ds),$$

where $PO(ns)$ is a Poisson random variable with parameter ns. Note for any $x>0$,

$$P[\frac{1}{n}PO(ns)\le x] \to \begin{cases} 1, & \text{if } x>s,\\ 0, & \text{if } x<s.\end{cases}$$

If F is continuous at x, then for any $\delta>0$

$$\sum_{j\le nx}\frac{(-1)^j}{j!}n^j\hat F^{(j)}(n)$$
$$= \int_0^{x-\delta} P[\frac{1}{n}PO(ns)\le x]F(ds) + \int_{x-\delta}^{x+\delta} P[\frac{1}{n}PO(ns)\le x]F(ds)$$
$$+ \int_{x+\delta}^\infty P[\frac{1}{n}PO(ns)\le x]F(ds).$$

By dominated convergence

$$\int_0^{x-\delta} P[\frac{1}{n}PO(ns)\le x]F(ds) \to F(x-\delta),$$
$$\int_{x+\delta}^\infty P[\frac{1}{n}PO(ns)\le x]F(ds) \to 0.$$

Therefore,

$$\liminf_{n\to\infty}\sum_{j\le nx}\frac{(-1)^j}{j!}n^j\hat F^{(j)}(n) \ge F(x-\delta) \to F(x)$$

as $\delta \to 0$ and

$$\limsup_{n\to\infty} \sum_{j\le nx} \frac{(-1)^j}{j!} n^j \hat{F}^{(j)}(n) \le F(x+\delta) + F(x-\delta, x+\delta] \to F(x),$$

as $\delta \to 0$. This shows the result.

7.7.23. We have

$$E(X_1^2) = \int_{-1}^{1} \frac{1}{2} x^2 dx = \frac{2}{6} = \frac{1}{3}.$$

The rest follows from the weak law of large numbers since

$$2^{-n}\lambda_n(B_{n,\delta} \cap I_n) = P[\sqrt{1/3} - \delta < \frac{\|\boldsymbol{X}\|_n}{\sqrt{n}} < \sqrt{1/3} + \delta] \to 1.$$

7.7.43. We have that the series defining Y converges from the fact that $\{\sum_{i=1}^{n} B_i/2^i, n \ge 1\}$ is non-decreasing in n and therefore has a limit. The limit must be finite since

$$\sum_{i=1}^{\infty} \frac{B_i}{2^i} \le \sum_{i=1}^{\infty} \frac{1}{2^i} = 1.$$

The range of Y is 0 (when all B_i's are 0) to 1 (when all B_i's are 1). We have

$$E(Y) = E\Big(\sum_{i=1}^{\infty} \frac{B_i}{2^i}\Big) = p\sum_{i=1}^{\infty} \frac{1}{2^i} = p,$$

and

$$\mathrm{Var}(Y) = \sum_{i=1}^{\infty} \frac{\mathrm{Var}(B_i)}{2^{2i}} = \frac{pq}{3}.$$

Let $x \in [0,1]$ be represented by its non-terminating dyadic expansion

$$x = .x_1x_2x_3\cdots = \sum_{i=1}^{\infty} \frac{x_i}{2^i},$$

where $x_i \in \{0,1\}$. Then Q_p concentrates on the following subset of $[0,1]$:

$$\Lambda_p := \{x \in [0,1] : \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = p\}.$$

Consequently, if $p \ne p'$, then

$$Q_p(\Lambda_p) = 1, \quad Q_p(\Lambda_{p'}) = 0.$$

(b) Denote a dyadic interval of length $1/2^{n+1}$ by

$$I(b_1,\ldots,b_n) := [.b_1\ldots b_n 0, .b_1\ldots b_n 1).$$

Then

$$\begin{aligned}P[Y \in I(b_1,\dots,b_n)] =& P[B_i = b_i,\ i = 1,\dots,n,\ B_{n+1} = 0]\\ =& p^{\sum_{i=1}^n b_i} q^{n-\sum_{i=1}^n b_i} q.\end{aligned}$$

So for $x \in [0,1]$, $x \in I(x_1,\dots,x_n)$ and therefore

$$Q_p(\{x\}) \leq P[Y \in I(x_1,\dots,x_n) = p^{\sum_{i=1}^n x_i} q^{n-\sum_{i=1}^n x_i} q \to 0,$$

as $n \to \infty$ and therefore $Q_p(\{x\}) = 0$ and $F_p(\cdot)$ is continuous.

To see that F_p is strictly continuous, note that if $x_1 < x_2$, then for big enough n, there is a dyadic interval I on the $1/2^n$ grid which is contained in $(x_1, x_2]$ and therefore

$$Q_p(x_1, x_2] \geq Q_p(I) > 0.$$

If $x \leq 1/2$,

$$\begin{aligned}P[\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq x] =& P[B_1 = 0, \sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq x]\\ =& qP[\sum_{i=2}^{\infty} \frac{B_i}{2^i} \leq x] = qP[\frac{1}{2}\sum_{i=1}^{\infty} \frac{B_{i+1}}{2^i} \leq x]\\ =& qP[\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq 2x],\end{aligned}$$

since $\{B_j, j \geq 2\} \stackrel{d}{=} \{B_j, n \geq 1\}$.

If $\frac{1}{2} \leq x \leq 1$, then

$$P[\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq x] = P[\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq x, B_1 = 0] + P[\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq x, B_1 = 1]$$

and since $B_1 = 0$ implies $Y \leq 1/2$ we get

$$\begin{aligned}=& q + pP[\frac{1}{2} + \sum_{i=2}^{\infty} \frac{B_i}{2^i} \leq x]\\ =& q + pP[\frac{1}{2} + \frac{1}{2}\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq x]\\ =& q + pP[\sum_{i=1}^{\infty} \frac{B_i}{2^i} \leq 2(x - \frac{1}{2})].\end{aligned}$$

7.7.45. By the strong law of large numbers, we have

$$\begin{aligned}\frac{-\log p_n(X_1,\ldots,X_n)}{n} =& \frac{-1}{n}\sum_{i=1}^{n} p_{X_i} \\ &\xrightarrow{\text{a.s.}} E(-\log p_{X_1}) = -\sum_{i=1}^{r} p_i \log p_i = H.\end{aligned}$$

7.7.46. For the Cauchy distribution, $E(|X_1|) = \infty$ so the conclusion for the sums follows by the strong law of large numbers.

Note as $x \to \infty$

$$P[X_1 > x] = \int_x^\infty \frac{1}{\pi(1+u^2)}du \sim \frac{1}{\pi x}$$

so for $x > 0$

$$nP[X_1 > nx/\pi] \sim n \cdot \frac{1}{\pi n x} \to x^{-1}.$$

Therefore, for $x > 0$

$$\begin{aligned}P[\bigvee_{i=1}^{n} X_i \le nx/\pi] =&(P[X_1 \le nx/\pi])^n = \left(1 - \frac{nP[X_1 > nx/\pi]}{n}\right)^n \\ &\to \exp\{-x^{-1}\}.\end{aligned}$$

Chapter 8 Solutions

8.8.2. (a) We have $X_n \xrightarrow{P} 0$ since $P[|X_n| > \epsilon] = \frac{1}{n} \to 0$.

(b) We have $X_n \Rightarrow 0$ since for a constant limit, convergence in probability is equivalent to convergence in distribution.

(c) We have

$$\sum_n P[X_n = n] = \sum_n \frac{1}{n} = \infty$$

and therefore

$$P\{[X_n = n] \text{ i.o. }\} = 1$$

and $\limsup_{n\to\infty} \frac{X_n}{n} = 1$. We conclude $\limsup_{n\to\infty} X_n = \infty$ almost surely.

Also $\liminf_{n\to\infty} X_n = 0$ since $X_n \xrightarrow{P} 0$ implies that for some subsequence $\{n_k\}$, we have $X_{n_k} \xrightarrow{\text{a.s.}} 0$. Thus $\liminf_{n\to\infty} X_n \le 0$. Since also $P[X_n \ge 0] = 1$ we have $\liminf_{n\to\infty} X_n = 0$.

8.8.3. Note $-\log U \stackrel{d}{=} E$ where $P[E > x] = e^{-x}$, $x > 0$. By the strong law of large numbers,

$$-\log \prod_{j=1}^n (U_j)^{\frac{1}{n}} = \frac{-\sum_{j=1}^n \log U_j}{n} \stackrel{d}{=} \frac{\sum_{i=1}^n E_i}{n} \to 1 \text{ a.s.}$$

By the central limit theorem

$$\sum_{j=1}^n \frac{-\log U_j - n}{\sqrt{n}} \Rightarrow N(0,1)$$

so

$$\sqrt{n}\left(\sum_{j=1}^n -\log U_j^{\frac{1}{n}} - 1\right) \Rightarrow (N(0,1).$$

Let $g(x) = e^{-x}$. By the delta method,

$$\begin{aligned}\sqrt{n}\left(\prod_1^n U_j^{\frac{1}{n}} - e^{-1}\right) &= \sqrt{n} g\left(\frac{1}{n}\sum_{j=1}^n(-\log U_j) - g(1)\right)\\ &\Rightarrow g'(1)N(0,1) = -e^{-1}N(0,1).\end{aligned}$$

8.8.4. (a) Suppose X_n has distribution F_n for $n \ge 0$. Suppose $X_n \Rightarrow X_0$. Given k and $\varepsilon < \frac{1}{2}$, we have $(k-\varepsilon, k+\varepsilon)$ is an interval of continuity of $F_0(x)$ and so

$$P[X_n = k] = P[X_n \in (k-\varepsilon, k+\varepsilon]] \to P[X_0 \in (k-\varepsilon, k+\varepsilon]] = P[X_0 = k].$$

Conversely, suppose $a < b$ are not integers. Given that $P[X_n = k] \to P[X_0 = k]$, we have

$$P[X_n \in (a,b]] = \sum_{k \in (a,b]} P[X_n = k] \to \sum_{k \in (a,b]} P[X_0 = k] = P[X_0 \in (a,b]],$$

so $P[X_n \in I] \to P[X_0 \in I]$ for intervals of continuity.

(b) Let μ be counting measure on the integers so that

$$\mu(A) = \# \text{ integers } k \in A = \sum_k 1_A(k).$$

Set $p_n(x) = \sum_k P[X_n = k] 1_{\{k\}}(x)$. Define

$$F_n(A) = P[X_n \in A] = \sum_{k \in A} P[X_n = k] = \int_A p_n(x)\mu(dx).$$

From (a), $F_n \Rightarrow F_0$ iff $P[X_n = k] \to P[X_0 = k]$ iff $p_n(x) \to p_0(x)$, for all integral x. But according to Scheffé's lemma, $p_n(x) \to p_0(x)$ implies $p_n \to p_0$ in $L_1(d\mu)$, that is

$$\sum_k |P[X_n = k] - P[X_0 = k]| = \int |p_n(x) - p_0(x)|\mu(dx) \to 0.$$

(c) From (a): $1_{A_n} \to 1_{A_0}$ iff

$$P[1_{A_n} = 1] = P(A_n) \to P[1_{A_0} = 1] = P(A_0).$$

(d) Suppose $x_n \to x_0$. Let f be bounded and continuous. Then

$$\int f dF_n = f(x_n) \to f(x_0) = \int f dF_0$$

and hence $F_n \Rightarrow F_0$ by the Portmanteau theorem.

Conversely, if $x_n \not\to x_0$, then there exists $\varepsilon > 0$, and there exists a subsequence $\{n'\}$ such that $|x_{n'} - x_0| > \varepsilon$. Define $f(x) = |x_0 - x| \wedge 1$ which is bounded and continuous. Then

$$\int f dF_{n'} = |x_0 - x_{n'}| \wedge 1 \geq \varepsilon \quad \text{and} \quad \int f dF_0 = |x_0 - x_0| = 0$$

so $\int f dF_n \not\to \int f dF_0$ and $F_n \not\Rightarrow F_0$.

(e) For f bounded and continuous

$$\begin{aligned} Ef(X_n) &= \frac{1}{2} f(1 - \frac{1}{n}) + \frac{1}{2} f(1 + \frac{1}{n}) \to \frac{1}{2} f(1) + \frac{1}{2} f(1) \\ &= f(1) = Ef(x) \end{aligned}$$

and therefore $X_n \Rightarrow X$. Define the mass functions

$$f_n(x) = \begin{cases} 0, & \text{if } x \notin \{1+\frac{1}{n}, 1-\frac{1}{n}\} \\ \frac{1}{2}, & \text{if } x = 1-\frac{1}{n}, \\ \frac{1}{2}, & \text{if } x = 1+\frac{1}{n} \end{cases} \qquad f_0(x) = \begin{cases} 1, & \text{if } x = 1, \\ 0, & \text{if } x \neq 1. \end{cases}$$

Then for every x, $f_n(x) \not\to f_0(x)$.

8.8.5. (a) On $[a,b]$, u_0 is uniformly continuous so given $\varepsilon > 0$, there exists $\delta > 0$ such that for $t_1, t_2 \in [a,b]$ we have

$$|t_1 - t_2| < \delta \text{ implies } |u_0(t_1) - u_0(t_2)| < \varepsilon. \tag{*}$$

Let $N_\delta(t) = \{s : |t-s| < \delta\}$ be a δ-neighborhood of t. Then

$$[a,b] \subset \bigcup_{t \in [a,b]} N_\delta(t)$$

and by compactness of $[a,b]$, there is a finite subcover

$$[a,b] \subset \bigcup_{i=1}^{k} N_\delta(t_i).$$

Without loss of generality we can suppose

$$t_0 = a \le t_1 < t_2 < \dots < t_k = b$$

and

$$\bigvee_{i=1}^{k} |t_i - t_{i-1}| < \frac{\delta}{2},$$

since if necessary, we can increase the subcover to achieve this. Pick n_0 so large that for $n \ge n_0$

$$\sup_{0 \le i \le n} |u_n(t_i) - u_0(t_i)| < \varepsilon. \tag{\#}$$

For $x \in [t_{i-1}, t_i]$, since u_n is nondecreasing, we get for $n \ge n_0$ that

$$\begin{aligned} |u_n(x) - u_0(x)| &\le |u_n(x) - u_n(t_i)| + |u_n(t_i) - u_0(t_i)| + |u_0(t_i) - u_0(X)| \\ &= A + B + C. \end{aligned}$$

For A we have

$$\begin{aligned} A &\le u_n(t_i) - u_n(t_{i-1}) && \text{(by monotonicity)} \\ &\le u_0(t_i) + \varepsilon - (u_0(t_{i-1}) - \varepsilon) && \text{(from \#)} \\ &= 2\varepsilon + u_0(t_i) - u_0(t_{i-1}) \le 3\varepsilon && \text{(from *).} \end{aligned}$$

We also get $B \le \varepsilon$ (from #) and $C \le \varepsilon$ (from *). We conclude that

$$\sup_{x\in[a,b]} |u_n(x) - u_0(x)| \le \bigvee_{i=1}^{k} \sup_{x\in[t_i - t_{i-1}]} |u_n(x) - u_0(x)| \le 5\varepsilon,$$

provided $n \ge n_0$.

(b) Given $\varepsilon > 0$, pick $M > 0$ such that

$$F_0(-M) \bigvee (1 - F_0(M)) \le \varepsilon. \tag{\#\#}$$

Then pick n_0 such that for $n \ge n_0$

$$|F_n(\pm M) - F_0(\pm M)| \le \varepsilon \text{ and } \sup_{[-M,M]} |F_n(x) - F_0(x)| \le \varepsilon. \tag{**}$$

Then we have

$$\begin{aligned}
\sup_{x\in\mathbb{R}} |F_n(x) - F_0(x)| \le & \sup_{x\le M} \bigvee \sup_{-M\le x\le M} \bigvee \sup_{x\ge M} \\
\le & (F_n(M) \bigvee F_0(M)) \bigvee \sup_{x\in[-M,M]} |F_n(x) - F_0(x)| \\
& \bigvee \Big[(1 - F_n(M)) \bigvee (1 - F_0(M)) \Big] \\
\le & ((F_0(M) + \varepsilon) \bigvee F_0(M)) \bigvee \varepsilon \\
& \bigvee (1 - F_0(M) + \varepsilon) \bigvee (1 - F_0(M)) \\
\le & 2\varepsilon \bigvee \varepsilon \bigvee 2\varepsilon = 2\varepsilon.
\end{aligned}$$

(c) We now verify the Glivenko-Cantelli lemma for this special case: Let $F_n(x,\omega) = \sum_{j=1}^{n} 1_{[X_j \le x]}(\omega)$. There exists $N_x \in \mathcal{B}$, such that $PN_x = 0$ and if $\omega \in N_x^c$ then $F_n(x,\omega) \to F(x)$. Let $\mathbb{Q}$ denote the rational numbers and we have that

$$\omega \in \bigcap_{x\in\mathbb{Q}} N_x^c \text{ implies } F_n(x,\omega) \to F(x).$$

Now $\Lambda := \bigcap_{x\in\mathbb{Q}} N_x^c \in \mathcal{B}$, and $P(\Lambda) = 1$ and for $\omega \in \Lambda$, $F_n(\cdot,\omega) \xrightarrow{w} F(\cdot)$. Hence by (b), for all $\omega \in \Lambda$:

$$\sup_{x\in\mathbb{R}} |F_n(x,\omega) - F(x)| \to 0.$$

8.8.6. (i) If $F(ax + b) = F(cx + d)$ then

$$\frac{F^{\leftarrow}(y) - b}{a} = \frac{F^{\leftarrow}(y) - d}{c}.$$

Proceed as in convergence to types theorem.

(ii) It is enough to show $F(Ax+B) = F(x)$ implies $A = 1, B = 0$. If $A = 1$, then

$$F(y) = F(y+B) = F(y+B+B) = F(y+2B) = \cdots = F(y+nB)$$

and letting $n \to \infty$ gives $F(y) = 1$, (if $B > 0$) or $= 0$ (if $B < 0$) for all y which contradicts the fact that F is non-degenerate. If $A \neq 1$, then

$$\begin{aligned} F(y + \frac{B}{1-A}) =& F(A(y + \frac{B}{1-A} + B) = F(Ay + \frac{AB+B-AB}{1-A}) \\ =& F(Ay + \frac{B}{1-A}). \end{aligned}$$

Define $G(y) = F(y + \frac{B}{1-A})$ and $G(y) = G(Ay)$. So iterating, we get $G(y) = G(A^n y)$. If $A > 1, y > 0$ we get $G(y) = G(\infty) = 1$ and if $y < 0, G(y) = G(-\infty) = 0$ so G is degenerate at 0. If $A < 1$ then $G(y) = G(0)$, forall y which also contradicts G being proper and non-degenerate.

8.8.8. We have

$$\begin{aligned} P[X_{\ell,n} \le x] =& P[\text{at least } \ell \text{ observations } \le x] \\ =& \sum_{k=\ell+1}^{n} \binom{n}{k} F^k(x)(\overline{F}(x))^{n-k} \end{aligned}$$

and thus the density is

$$f_{X_{\ell,n}}(x) = F^{\ell-1}(x) f(x)(1-F(x))^{n-\ell} \frac{n!}{(\ell-1)!(n-\ell)!}.$$

Since $F(x) = 1 - e^{-x}$ for $x > 0$, we get

$$\begin{aligned} \frac{1}{n} f_{X_{\ell,n}}(\frac{x}{n}) =& \frac{1}{n}(1-e^{-\frac{x}{n}})^{\ell-1} e^{-\frac{x}{n}} (e^{-\frac{x}{n}})^{n-\ell} \ \frac{n!}{(\ell-1)!(n-\ell)!} \\ \sim& \frac{1}{n}(\frac{x}{n})^{\ell-1}(1+o(1)) e^{-\frac{x}{n}} e^{-x} \ \frac{n!}{(\ell-1)!(n-\ell)!} \\ \sim& \frac{x^{\ell-1}e^{-x}}{(\ell-1)!} \ \frac{n(n-1)\dots(n-\ell+1)}{n^\ell} \\ \sim& \frac{x^{\ell-1}e^{-x}}{(\ell-1)!}. \end{aligned}$$

8.8.10. We have for any $\delta > 0$ that

$$P[|X_n - X| > \delta] \ge P[Y = 0, X = 1] = \frac{1}{4}.$$

This does not converge to 0. Note $X_n \Rightarrow X$ since $X_n \stackrel{d}{=} X$.

8.8.13. If $\mu_n \to \mu_0$, and $\sigma_n \to \sigma_0$ then the densities converge

$$n(\mu_n, \sigma_n, x) \to n(\mu_0, \sigma_0, x)$$

as $n \to \infty$ and Scheffe lemma applies.

8.8.23. (i) We use the delta method with $g(x) = x^2$ to get

$$\begin{aligned}\sqrt{n}(g(\bar{X}_n) - g(\mu)) =& \frac{g(\bar{X}_n) - g(\mu)}{\bar{X}_n - g(\mu)} \cdot \sqrt{n}(\bar{X}_n - g(\mu)) \\ \Rightarrow& g'(\mu)N(0,\sigma^2) = 2\mu\sigma N.\end{aligned}$$

(ii) Use $g(x) = e^x$ so the limit is $g'(\mu)N(0,\sigma^2) \stackrel{d}{=} e^{\mu}\sigma N(0,1)$.

8.8.34. Observe that if $P[E_1 > x] = e^{-x}$, for $x > 0$, then with $g(x) = x^{1/\alpha}$ we have

$$P[g(E_1) > x] = P[E_1 > g^{\leftarrow}(x)] = e^{-x^{\alpha}}, \quad x > 0,$$

so that $\{W_n, n \geq 1\} \stackrel{d}{=} \{g(E_n), n \geq 1$ and because $g(\cdot)$ is non-decreasing

$$\bigvee_{i=1}^{n} W_i \stackrel{d}{=} \bigvee_{i=1}^{n} g(E_i) = g(\bigvee_{i=1}^{n} E_i).$$

Write

$$Y_n := \bigvee_{i=1}^{n} E_i - \log n \Rightarrow Y.$$

By the Baby Skorohod theorem, there exists $Y_n^{\#}$ and $Y^{\#}$ such that $Y_n^{\#} \stackrel{d}{=} Y_n$ and $Y^{\#} \stackrel{d}{=} Y$ and $Y_n^{\#} \stackrel{\text{a.s.}}{\to} Y^{\#}$. So we have

$$\begin{aligned}\frac{\bigvee_{i=1}^{n} W_i - g(\log n)}{g'(\log n)} \stackrel{d}{=}& \frac{g(Y_n^{\#} + \log n) - g(\log n)}{g'(\log n)} \\ =& \frac{\int_{\log n}^{Y_n^{\#}+\log n} g'(s)ds}{g'(\log n)}\end{aligned}$$

which by the mean value theorem is

$$= \frac{g'(\zeta_n^{\#})}{g'(\log n)} Y_n^{\#},$$

where $\zeta_n^{\#}$ is between $\log n$ and $Y_n^{\#} + \log n$. Since $Y_n^{\#} \stackrel{\text{a.s.}}{\to} Y^{\#}$, we have $\zeta_n^{\#}/\log n \stackrel{\text{a.s.}}{\to} 1$ from which follows

$$\frac{g'(\zeta_n^{\#})}{g'(\log n)} = \Big(\frac{\zeta_n^{\#}}{\log n}\Big)^{1/\alpha - 1} \stackrel{\text{a.s.}}{\to} 1.$$

Thus

$$\frac{\bigvee_{i=1}^{n} W_i - g(\log n)}{g'(\log n)} \stackrel{d}{=} \frac{g'(\zeta_n^{\#})}{g'(\log n)} Y_n^{\#} \stackrel{\text{a.s.}}{\to} 1 \cdot Y^{\#} \stackrel{d}{=} Y.$$

Chapter 9 Solutions

9.9.2. (a) We have

$$E(e^{itX_n}) = e^{\lambda_n(e^{it}-1)}$$

and therefore $X_n \Rightarrow X_0$ iff $\lambda_n \to \lambda_0$.

(b) If $\mu_n \to \mu_0$, and $\sigma_n \to \sigma_0$ then

$$\phi_{X_n}(t) = e^{it\mu_n}e^{-\sigma_n^2 t^2/2} \to e^{it\mu_0 - \sigma^2 t^2/2}$$

so

$$X_n \sim N(\mu_n, \sigma_n^2) \Rightarrow X_0 \sim N(\mu_0, X_0^2).$$

Conversely: Suppose $X_n \Rightarrow X_0$. Let $\{X_n'\}$ be iid and independent of $\{X_n\}$ and have the same marginal distribution: $X_n' \stackrel{d}{=} X_n$. Then $X_n - X_n' \Rightarrow X_0 - X_0'$. So

$$\phi_{X_n - X_n'}(t) = e^{-2\sigma_n^2 t^2/2} \to e^{-2\sigma_0^2 t^2/2} = \phi_{X_0 - X_0'}(t),$$

and therefore $\sigma_n \to \sigma_0$. But if $\sigma_n \to \sigma_0$, then from convergence of the chf's we get $\mu_n \to \mu_0$.

9.9.3. We have

$$EX_k^2 = 1 - \frac{1}{k^2} + k^2 \cdot \frac{1}{k^2} = 2 - \frac{1}{k^2} \to 2,$$

and therefore,

$$\mathrm{Var}\left(\frac{S_n}{\sqrt{n}}\right) = \frac{1}{n}\sum_{k=1}^{n}(2 - \frac{1}{k^2}) \to 2.$$

Next, let

$$X_k^* = \begin{cases} 1, & \text{if } X_k = 1 \text{ or } k \\ -1, & \text{if } X_k = -1 \text{ or } -k. \end{cases}$$

Thus $\{X_k^*\}$ are Bernoulli random variables with

$$P[X_k^* = 1] = \frac{1}{2} = P[X_k^* = -1]$$

and $S_n^*/\sqrt{n} \Rightarrow N(0,1)$. Let

$$\begin{aligned} m_0 &= \sup\{m : \sum_{k=1}^{m} k \le \epsilon\sqrt{n}\} = \sup\{m : \frac{m}{2}(1+m) \le \epsilon\sqrt{n}\} \\ &\ge \sup\{m : (m+1)^2 \le 2\epsilon\sqrt{n}\} = \sqrt{2\epsilon\sqrt{n}} - 1\}. \end{aligned}$$

If $|S_n^* - S_n| > \epsilon\sqrt{n}$, then it means that there exists $i \geq m_0$ such that $X_i \neq X_i^*$, since if $i < m_0$ we would not get a big enough contribution to the difference. Thus

$$\begin{aligned} P[|S_n - S_n^*| > \epsilon\sqrt{n}] \leq& P\{ \bigcup_{m_0 \leq k \leq n} [X_k \neq X_k^*]\} \\ \leq& \sum_{k=m_0}^{n} \frac{1}{k^2} \leq \sum_{k=\sqrt{2\epsilon\sqrt{n}}-1}^{n} \frac{1}{k^2} \\ \sim& (\sqrt{2\epsilon} n^{\frac{1}{4}})^{-1} - n^{-1} \to 0. \end{aligned}$$

It follows that

$$\frac{S_n}{\sqrt{n}} - \frac{S_n^*}{\sqrt{n}} \xrightarrow{P} 0$$

and by Slutsky's lemma $\frac{S_n}{\sqrt{n}} \Rightarrow N(0,1)$.

9.9.4. (a) We have

$$EU_k^2 = \frac{a_k^2}{3}, \quad E|U_k|^3 = a_k^3/4$$

and the Liapunov condition becomes

$$\begin{aligned} \sum_{k=1}^{n} \frac{E|U_k|^3}{s_n^3} =& \text{const } \frac{\sum_{k=1}^{n} a_k^3}{(\sum_{k=1}^n a_k^2)^{\frac{3}{2}}} \leq M\text{const} \frac{\sum_{k=1}^n a_k^2}{(\sum_{k=1}^n a_k^2)^{\frac{3}{2}}} \\ =& \frac{M \text{ const}}{(\sum_{k=1}^n a_k^2)^{\frac{1}{2}}} \to 0. \end{aligned}$$

(b) We have $s_n^2 = \frac{1}{3}\sum_1^n a_k^2 \uparrow s_\infty < \infty$ so

$$\frac{1}{s_n^2}\sum_1^n EU_k^2 1_{[|\frac{U_k}{s_n}|>\epsilon]} \geq \frac{1}{s_\infty^2}\sum_1^n EU_k^2 1_{[|U_k|>\epsilon s_\infty]} \not\to 0.$$

9.9.5. We know that

$$\phi_{X_n}(t) \to \phi_{X_0}(t) \text{ and } \phi_{Y_n}(t) \to \phi_{Y_0}(t)$$

so if X_n is independent of Y_n we have

$$\phi_{X_n+Y_n}(t) = \phi_{X_n}(t)\phi_{Y_n}(t) \to \phi_{X_0}(t)\phi_{Y_0}(t) = \phi_{X_0+Y_0}(t).$$

9.9.6. (a) We have

$$\phi_n(t) = \int_{-n}^{n} e^{itx}\frac{1}{n}dx = \frac{2\cos tn}{itn} \to 0.$$

(b) This follows from the cosine funtion being bounded by 1.

(c) No. This does not contradict the continuity then since the limit does not satisfy $\phi(0) = 1$.

9.9.7. If

$$f(x) = x^{-3}, \quad |x| > 1,$$

then $E(X_n) = 0$, and $E(X_n^2) = \infty$. Set

$$Y_n = X_n 1_{[|X_n| \le \sqrt{n}\log n]}.$$

Then

$$\begin{aligned} E(Y_n^3) =& E|X_n|^3 1_{[|X_n| \le \sqrt{n}\log n]} = 2\int_1^{\sqrt{n}\log n} x^3 x^{-3} dx \\ =& \sqrt{n}\log n - 1 \sim \sqrt{n}\log n. \end{aligned}$$

Also, we have

$$\begin{aligned} E(Y_n^2) =& E(X_n^2 1_{[|X_n| \le \sqrt{n}\log n]}) = 2\int_1^{\sqrt{n}\log n} x^2 x^{-3} dx \\ =& 2\log(\sqrt{n}\log n) \sim \log n. \end{aligned}$$

Therefore,

$$s_n^2 = \sum_1^n E(Y_j^2 \sim \int_1^n \log x dx \sim n\log n$$

and

$$\frac{\sum_1^n E|Y_n|^3}{s_n^3} \sim \frac{\int_1^n \sqrt{x}\log x dx}{(n\log n)^{3/2}} \sim \frac{n^{\frac{3}{2}}\log n}{n^{\frac{3}{2}}(\log n)^{\frac{3}{2}}} \sim \frac{1}{(\log n)^{\frac{1}{2}}} \to 0.$$

So we see that the Liapunov condition holds for $\{Y_n\}$ and consequently $\sum_{j=1}^n Y_j/s_n \Rightarrow N$ and by the convergence to types theorem

$$\sum_{j=1}^n Y_j/\sqrt{n\log n} \Rightarrow N.$$

Lastly observe that

$$\begin{aligned} \sum_n P[X_n \ne Y_n] =& \sum_n P[|X_n| > \sqrt{n}\log n] = \sum_n 2\int_{\sqrt{n}\log n}^{\infty} x^{-3} dx \\ =& \sum_n \frac{1}{n(\log n)^2} < \infty. \end{aligned}$$

The demonstration can now be completed with the Equivalence Proposition 7.1.1 since we have

$$\frac{\sum_{j=1}^{n} X_j}{s_n} - \frac{\sum_{j=1}^{n} Y_j}{s_n} \xrightarrow{\text{a.s.}} 0.$$

9.9.8. (a) Observe that if $\{X_n\}$ is iid with Poisson distribution with parameter 1, then $S_n = \sum_{1}^{n} X_i$ is Poisson distributed with parameter n. Therefore, we have that

$$\begin{aligned} E\Big(\frac{S_n - n}{\sqrt{n}}\Big)^- &= \sum_{k=0}^{n} \Big(\frac{n-k}{\sqrt{n}}\Big) P[S_n = k] = \sum_{k=0}^{n} \frac{n-k}{\sqrt{n}} \frac{e^{-n} n^k}{k!} \\ &= \frac{e^{-n}}{\sqrt{n}} \sum_{k=0}^{n} (n-k) \frac{n^k}{k!} = \frac{e^{-n}}{\sqrt{n}} \left(\sum_{k=0}^{n} \Big(\frac{n^{k+1}}{k!} - \frac{n^k}{(k-1)!} \Big) \right) \end{aligned}$$

and because this last sum telescopes, we get

$$= \frac{e^{-n} n^{n+1}}{\sqrt{n} n!} = \frac{e^{-n} n^{n+\frac{1}{2}}}{n!}.$$

(b) From the central limit theorem,

$$\frac{S_n - n}{\sqrt{n}} \Rightarrow N,$$

where N is a standard normal random variable, since $E(X_n) = 1$, and $\mathrm{Var}(X_n) = 1$. From the continuous mapping theorm

$$\left(\frac{S_n - n}{\sqrt{n}}\right)^- \Rightarrow N^-.$$

(c) We have

$$\left(\frac{S_n - n}{\sqrt{n}}\right)^- \leq \left|\frac{S_n - n}{\sqrt{n}}\right|$$

and therefore,

$$\left(\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right)^2 \leq \left(\frac{S_n - n}{\sqrt{n}}\right)^2,$$

so that

$$\sup_n E\left(\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right)^2 \leq \sup_n E\left(\frac{S_n - n}{\sqrt{n}}\right)^2 = 1.$$

So from the Crystal Ball Condition (6.13) we get that

$$\left(\frac{S_n - n}{\sqrt{n}}\right)^- \text{ is ui.}$$

Therefore the mean of the displayed random variable converges to $E(N^-)$.

(d) We compute

$$E(N^-) = E(N^+) = \int_0^\infty x\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx$$

and making the change of variable $y = x^2/2$ so that $dy = xdx$ we get

$$= \int_0^\infty \frac{e^{-y}dy}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}.$$

We conclude

$$\frac{n^{n+\frac{1}{2}}e^{-n}}{n!} \to \sqrt{2\pi}$$

or

$$n! \sim \frac{1}{\sqrt{2\pi}}e^{-n}n^{n+1/2}.$$

9.9.9. (a) We have

$$\phi(t) = \int_0^\infty e^{itx}e^{-x}dx = \frac{1}{1-it}$$

Also, note that $\frac{1}{1+it} = \phi(-t)$ which is the characteristic function of $-X$.

(b) The chf of X_1 is $\frac{e^{it}+e^{-it}}{2} = \cos t$.

(c) Observe that $(\cos t)^{17}$ is the chf of $X_1 + \cdots + X_{17}$ where $\{X_n\}$ are iid Bernoulli.

(d) We have

$$\begin{aligned}\frac{2}{\pi}\int_0^\infty \frac{1 - \operatorname{Re}\phi(t)}{t^2}dt =& \frac{2}{\pi}\int_0^\infty \frac{1 - E(\cos tX)}{t^2}dt\\ =& \frac{2}{\pi}\int_0^\infty E\left(\frac{1-\cos tX}{t^2}\right)dt\end{aligned}$$

and applying Fubini's theorem, and then the fact that cos is an even function we get

$$\begin{aligned}=& \frac{2}{\pi}E\int_0^\infty \left(\frac{1-\cos tX}{t^2}\right)dt\\ =& \frac{2}{\pi}E\int_0^\infty \left(\frac{1-\cos t|X|}{t^2}\right)dt\end{aligned}$$

and with the change of variable $s = t|X|$, we get

$$=2\int_0^\infty \left(\frac{1-\cos(s)}{\pi s^2}\right)dsE|X| = 2E|X|,$$

since $(1-\cos(s))/(\pi s^2)$ is a density.

9.9.10. We have

$$E(e^{it(Y_s-s)/\sqrt{s}}) = \exp\{s[e^{it/\sqrt{s}} - 1 - it/\sqrt{s}]\}$$

so it suffices to show

$$\left|s[e^{it/\sqrt{s}} - 1 - it/\sqrt{s}\,] + t^{\frac{2}{2}}\right| \to 0.$$

However

$$\begin{aligned}\left|s[e^{it/\sqrt{s}} - 1 - \frac{it}{\sqrt{s}}] + \frac{1}{2}t^2\right| =& \left|s[e^{it/\sqrt{s}} - 1 - \frac{it}{\sqrt{s}} - \frac{1}{2}\left(it/\sqrt{s}\right)^2]\right| \\ \leq& s\left|[e^{it/\sqrt{s}} - 1 - \frac{it}{\sqrt{s}} - \frac{1}{2}\left(it/\sqrt{s}\right)^2]\right| \\ \leq& \frac{s}{3}\left|\frac{it}{\sqrt{s}}\right|^3 = O(\frac{1}{s^{1/2}}) \to 0,\end{aligned}$$

as $s \to \infty$.

9.9.12. Let $X_j' = X_j 1_{[|X_j| \leq ta_n]}$ so that

$$P\left[\sum_{j=1}^{n}\frac{X_j'}{a_n} \neq \sum_{j=1}^{n}\frac{X_j}{a_n}\right] \leq P\left[\bigcup_{j=}^{n}[|X_j| > ta_n]\right] \leq \sum_{j=1}^{n}P[|X_j| > ta_n] \to 0.$$

So

$$\frac{S_n'}{a_n} - \frac{S_n}{a_n} \xrightarrow{P} 0$$

and it suffices to show $\frac{S_n'}{a_n} \Rightarrow N$ by Slutsky's theorem. Mimic the standard proof: The statement

$$Ee^{itS_n'/a_n} = \prod_{1}^{n}Ee^{itX_j'/a_n} \to e^{-t^2/2} \tag{1}$$

follows from

$$\sum_{j=1}^{n}E\left(e^{itX_j'/a_n}\right) - 1 + t^2/2 \to 0 \tag{2}$$

since

$$\exp\{\sum_{j=1}^{n}E\left(e^{itX_j'/a_n}\right) - 1\} - \prod_{j=1}^{n}Ee^{itX_j'/a_n} \to 0. \tag{3}$$

The reason for (3) is that by the product comparison lemma, the difference in (3) is

$$
\begin{aligned}
|3| \leq & \sum_{j=1}^{n} \left| \exp\{Ee^{itX_j'/a_n} - 1\} - Ee^{itX_j'/a_n} \right| \\
= & \sum_{j=1}^{n} \left| \exp\{Ee^{itX_j'/a_n} - 1\} - 1 - E(e^{itX_j'/a_n} - 1) \right| \\
= & \sum_{j=1}^{n} |e^{z_j} - 1 - z_j|,
\end{aligned}
$$

where $z_j = Ee^{itX_j'/a_n} - 1$. Now

$$
|z_j| = |Ee^{itX_j'/a_n} - 1| = |Ee^{itX_j'/a_n} - 1 - itX_j'/a_n|,
$$

(since $EX_j' = 0$ due to symmetry of $P[X_j \leq x]$)

$$
\begin{aligned}
\leq & E\left(\frac{1}{2} t^2 \frac{(X_j')^2}{a_n^2}\right) = \frac{t^2}{2} E\left(\frac{X_j}{a_n}\right)^2 1_{\left[\frac{|X_j|}{a_n} \leq t\right]} \qquad (*) \\
\leq & \frac{t^2}{2} \bigvee_{j=1}^{n} E\left(\frac{X_j}{a_n}\right)^2 1_{\left[\frac{|X_j|}{a_n} \leq t\right]} \to 0
\end{aligned}
$$

as $n \to \infty$. The reason for the convergence to 0 is that for any $\varepsilon > 0$

$$
\begin{aligned}
E\left(\frac{X_j}{a_n}\right)^2 1_{\left[\frac{|X_j|}{a_n} \leq t\right]} \leq & E\left(\frac{X_j}{a_n}\right)^2 1_{\left[\frac{|X_j|}{a_n} \leq \varepsilon\right]} + E\left(\frac{X_j}{a_n}\right)^2 1_{\left[\varepsilon < \frac{|X_j|}{a_n} \leq t\right]} \\
\leq & \varepsilon^2 + t^2 \sum_{j=1}^{n} P\left[\left|\frac{X_j}{a_n}\right| > \varepsilon\right] \to \varepsilon^2 + 0 = \varepsilon^2.
\end{aligned}
$$

We conclude that as $n \to \infty$, $|z_j| \to 0$ uniformly in j. It follows that for any $\delta > 0$ and n sufficiently large,

$$
|3| \leq \sum_{j=1}^{n} |e^{z_j} - 1 - z_j| \leq \delta \sum_{j=1}^{n} |z_j| \leq \frac{\delta t^2}{2} \sum_{j=1}^{n} E\left(\frac{(X_j')^2}{a_n^2}\right)
$$

(by $*$)

$$
= \frac{\delta t^2}{2} \sum_{j=1}^{n} E\left(\frac{X_j}{a_n} 1_{[|X_j| \leq a_n t]}\right) \to \frac{\delta t^2}{2}.
$$

Since δ is arbitrary $|3| \to 0$.

It remains to show (2) or what is the same thing

$$\sum_{j=1}^{n} E\left(e^{itX'_j/a_n} - 1 - \frac{\left(itX'_j/a_n\right)^2}{2}\right) \to 0, \qquad (2')$$

since

$$\sum_{j=1}^{n} E\left(\frac{X'_j}{a_n}\right)^2 = \sum_{1}^{n} E\left(\frac{X_j}{a_n}\right)^2 1_{\left[\left|\frac{X_j}{a_n}\right| \le t\right]} \to 1.$$

Now

$$|2'| \le \sum_{j=1}^{n} \frac{|t|^3}{3!} E\left(\left|\frac{X_j}{a_n}\right|^3 1_{\left[\left|\frac{X_j}{a_n}\right| \le t\right]}\right)$$

and for $\epsilon < t$ this is

$$= \sum_{j=1}^{n} \frac{|t|^3}{3!}\left(E\left|\frac{X_j}{a_n}\right|^3 1_{\left[\left|\frac{X_j}{a_n}\right| \le \epsilon\right]} + E\left|\frac{X_j}{a_n}\right|^3 1_{\left[\epsilon < \left|\frac{X_j}{a_n}\right| t\right]}\right)$$

$$\le \frac{\varepsilon|t|^3}{3!}\left(\sum_{j=1}^{n} E\left|\frac{X_j}{a_n}\right|^2 1_{\left[\left|\frac{X_j}{a_n}\right| \le \epsilon\right]}\right) + \frac{|t|^3}{3!}|t|^3 \sum_{j=1}^{n} P\left[\left|\frac{X_j}{a_n}\right| > \varepsilon\right]$$

and since the first sum converges to 1and the second converges to 0, we get

$$\to \frac{\varepsilon|t|^3}{3!} + 0.$$

Now $\varepsilon > 0$ is arbitrary so $2' \to 0$ as $n \to \infty$.

Breinigsville, PA USA
04 September 2009
223551BV00003B/16/A

http://www.springer.com/978-0-8176-4055-2

A Probability Path
Resnick, S.
1998, 230 p. 10 illus., Hardcover
ISBN: 978-0-8176-4055-2
A Birkhäuser Boston product