A Collection of Exercises in Advanced Probability Theory

The Solutions Manual of All Even-Numbered Exercises from "A First Look at Rigorous Probability Theory" (Second Edition, World Scientific, 2006)

Mohsen Soltanifar
PhD Candidate in Biostatistics, Dalla Lana School of Public Health
University of Toronto, Canada

Longhai Li
Professor of Statistics, Department of Mathematics and Statistics
University of Saskatchewan, Canada

Jeffrey S. Rosenthal
Professor of Statistics, Department of Statistical Sciences
University of Toronto, Canada

(Published July 2010; Updated January 2020; Corrections by Xinyu Huo August 2023)

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapor 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Web site: www.WorldScientific.com

A Collection of Exercises in Advanced Probability Theory

Preface

I am very pleased that, thanks to the hard work of Mohsen Soltanifar and Longhai Li, this solutions manual for my book[1] is now available. I hope readers will find these solutions helpful as you struggle with learning the foundations of measure-theoretic probability. Of course, you will learn best if you first attempt to solve the exercises on your own, and only consult this manual when you are really stuck (or to check your solution after you think you have it right).

For course instructors, I hope that these solutions will assist you in teaching students, by offering them some extra guidance and information.

My book has been widely used for self-study, in addition to its use as a course textbook, allowing a variety of students and professionals to learn the foundations of measure-theoretic probability theory on their own time. Many self-study students have written to me requesting solutions to help assess their progress, so I am pleased that this manual will fill that need as well.

Solutions manuals always present a dilemma: providing solutions can be very helpful to students and self-studiers, but make it difficult for course instructors to assign exercises from the book for course credit. To balance these competing demands, we considered maintaining a confidential "instructors and self-study students only" solutions manual, but decided that would be problematic and ultimately infeasible. Instead, we settled on the compromise of providing a publicly-available solutions manual, but to even-numbered exercises only. In this way, it is hoped that readers can use the even-numbered exercise solutions to learn and assess their progress, while instructors can still assign odd-numbered exercises for course credit as desired.

Of course, this solutions manual may well contain errors, perhaps significant ones. If you find some, then please e-mail me and I will try to correct them promptly. (I also maintain an errata list for the book itself, on my web site, and will add book corrections there.)

Happy studying!

Jeffrey S. Rosenthal, Toronto, Canada, 2010
(`jeff@math.toronto.edu` / `www.probability.ca`)

2020 update:

Some typos and incorrect or incomplete solutions have been fixed: eleven typos related to Exercises 2.6.4, 3.6.14, 3.6.18, 9.5.12, 10.3.2, 14.4.12(a), and 15.6.8; two partial solutions related to Exercises 2.7.14(a) and 8.5.20; and six incorrect solutions related to Exercises 2.7.22(a), 3.6.6(a), 4.5.10, 9.5.14(a), 13.4.6, and 13.4.10. Thank to Danny Cao, Felix Pahl and Byron Schmuland for pointing out the erroneous solutions.

In addition, the solutions to the even numbered Appendix Exercises A.3.2, A.3.8, A.4.4, and A.4.6 have been added, and two distinct solutions are now offered for Exercises 3.6.12, 4.5.10, 5.5.6, 11.5.6, and 12.3.4 to help the readers better grasp the key concepts and results.

For chapters 7, 8, 14, and 15, the reader may also wish to consult the related exercises in the new textbook *A First Look at Stochastic Processes* (J.S. Rosenthal, World Scientific Publishing, 2020).

2023 update:

Led by Xinyu Huo, a number of additional corrections and clarifications were made. Thanks to David Scott for pointing out many of them.

[1] J.S. Rosenthal, *A First Look at Rigourous Probability Theory*, 2nd ed. World Scientific Publishing, Singapore, 2006. 219 pages. ISBN: 981-270-371-5 / 981-270-371-3 (paperback).

Contents

1 The need for measure theory **1**

2 Probability triples **3**

3 Further probabilistic foundations **9**

4 Expected values **15**

5 Inequalities and convergence **20**

6 Distributions of random variables **23**

7 Stochastic processes and gambling games **25**

8 Discrete Markov chains **29**

9 More probability theorems **36**

10 Weak convergence **40**

11 Characteristic functions **45**

12 Decompositions of probability laws **49**

13 Conditional probability and expectation **52**

14 Martingales **56**

15 General stochastic processes **64**

A Mathematical Background **77**

Chapter 1

The need for measure theory

Exercise 1.3.2. Suppose $\Omega = \{1, 2, 3\}$ and $\mathcal{F}$ is a collection of all subsets of Ω. Find (with proof) necessary and sufficient conditions on the real numbers x, y, and z such that there exists a countably additive probability measure P on $\mathcal{F}$, with $x = P\{1, 2\}, y = P\{2, 3\}$, and $z = P\{1, 3\}$.

Solution. The necessary and sufficient conditions are: $0 \leq x \leq 1$, $0 \leq y \leq 1$, $0 \leq z \leq 1$, and $x + y + z = 2$.

To prove necessity, let P be a probability measure on Ω. Then, for

$$x = P\{1, 2\} = P\{1\} + P\{2\},$$

$$y = P\{2, 3\} = P\{2\} + P\{3\},$$

and

$$z = P\{1, 3\} = P\{1\} + P\{3\}$$

we have by definition that $0 \leq x \leq 1$, $0 \leq y \leq 1$, and $0 \leq z \leq 1$, and furthermore we compute that

$$x + y + z = 2(P\{1\} + P\{2\} + P\{3\}) = 2P(\Omega) = 2,$$

thus proving the necessity.

Conversely, assume that $0 \leq x \leq 1$, $0 \leq y \leq 1$, $0 \leq z \leq 1$, and $x + y + z = 2$. Then, define the desired countably additive probability measure P as follows:

$$\begin{aligned}
P(\phi) &= 0, \\
P\{1\} &= 1 - y, \\
P\{2\} &= 1 - z, \\
P\{3\} &= 1 - x, \\
P\{1, 2\} &= x, \\
P\{1, 3\} &= z, \\
P\{2, 3\} &= y, \\
P\{1, 2, 3\} &= 1.
\end{aligned}$$

It is easily checked directly that for any two disjoint sets $A, B \subseteq \Omega$, we have

$$P(A \cup B) = P(A) + P(B)$$

For example, if $A = \{1\}$ and $B = \{2\}$, then since $x + y + z = 2$, $P(A \cup B) = P\{1, 2\} = x$ while $P(A) + P(B) = P\{1\} + P\{2\} = (1 - y) + (1 - z) = 2 - y - z = (x + y + z) - y - z = x = P(A \cup B)$. Hence, P is the desired probability measure, proving the sufficiency.$\Box$

Exercise 1.3.4. Suppose that $\Omega = \mathbb{N}$, and P is defined for all $A \subseteq \Omega$ by $P(A) = |A|$ if A is finite (where $|A|$ is the number of elements in the subset A), and $P(A) = \infty$ if A is infinite. This P is of course not a probability measure(in fact it is counting measure), however we can still ask the following. (By convention, $\infty + \infty = \infty$.)
(a) Is P finitely additive?
(b) Is P countably additive?

Solution.(a) Yes. Let $A, B \subseteq \Omega$ be disjoint. We consider two different cases.

Case 1: At least one of A or B is infinite. Then $A \cup B$ is infinite. Consequently, $P(A \cup B)$ and at least one of $P(A)$ or $P(B)$ will be infinite. Hence, $P(A \cup B) = \infty$ and $P(A) + P(B) = \infty$, implying $P(A \cup B) = \infty = P(A) + P(B)$.

Case 2: Both of A and B are finite. Then $P(A \cup B) = |A \cup B| = |A| + |B| = P(A) + P(B)$.

Accordingly, P is finitely additive.

(b) Yes. Let $A_1, A_2, \cdots$ be a sequence of disjoint subsets of Ω. We consider two different cases.

Case 1: At least one of A_n's is infinite. Then $\cup_{n=1}^{\infty} A_n$ is infinite. Consequently, $P(\cup_{n=1}^{\infty} A_n)$ and at least one of $P(A_n)$'s will be infinite. Hence, $P(\cup_{n=1}^{\infty} A_n) = \infty$ and $\sum_{n=1}^{\infty} P(A_n) = \infty$, implying $P(\cup_{n=1}^{\infty} A_n) = \infty = \sum_{n=1}^{\infty} P(A_n)$.

Case 2: All of A_n's are finite. Then depending on finiteness of $\cup_{n=1}^{\infty} A_n$ we consider two cases. First, let $\cup_{n=1}^{\infty} A_n$ be infinite, then, $P(\cup_{n=1}^{\infty} A_n) = \infty = \sum_{n=1}^{\infty} |A_n| = \sum_{n=1}^{\infty} P(A_n)$. Second, let $\cup_{n=1}^{\infty} A_n$ be finite, then,

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

Accordingly, P is countably additive.$\Box$

Chapter 2

Probability triples

Exercise 2.7.2. Let $\Omega = \{1, 2, 3, 4\}$, and let $\mathcal{J} = \{\{1\}, \{2\}\}$. Describe explicitly the σ-algebra $\sigma(\mathcal{J})$ generated by $\mathcal{J}$.

Solution.

$$\sigma(\mathcal{J}) = \{\phi, \{1\}, \{2\}, \{1, 2\}, \{3, 4\}, \{1, 3, 4\}, \{2, 3, 4\}, \Omega\}.$$

□

Exercise 2.7.4. Let $\mathcal{F}_1, \mathcal{F}_2, \cdots$ be a sequence of collections of subsets of Ω, such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for each n.
(a) Suppose that each $\mathcal{F}_i$ is an algebra. Prove that $\cup_{i=1}^{\infty} \mathcal{F}_i$is also an algebra.
(b) Suppose that each $\mathcal{F}_i$ is a σ-algebra. Show (by counterexample) that $\cup_{i=1}^{\infty} \mathcal{F}_i$ might not be a σ-algebra.

Solution.(a) First, since $\phi, \Omega \in \mathcal{F}_1$ and$\mathcal{F}_1 \subseteq \cup_{i=1}^{\infty} \mathcal{F}_i$, we have $\phi, \Omega \in \cup_{i=1}^{\infty} \mathcal{F}_i$. Second, let $A \in \cup_{i=1}^{\infty} \mathcal{F}_i$, then $A \in \mathcal{F}_i$ for some i. On the other hand, $A^c \in \mathcal{F}_i$ and $\mathcal{F}_i \subseteq \cup_{i=1}^{\infty} \mathcal{F}_i$, implying $A^c \in \cup_{i=1}^{\infty} \mathcal{F}_i$. Third, let $A, B \in \cup_{i=1}^{\infty} \mathcal{F}_i$, then $A \in \mathcal{F}_i$ and $B \in \mathcal{F}_j$, for some i, j. However, $A, B \in \mathcal{F}_{\max(i,j)}$ yielding $A \cup B \in \mathcal{F}_{\max(i,j)}$. On the other had, $\mathcal{F}_{\max(i,j)} \subseteq \cup_{i=1}^{\infty} \mathcal{F}_i$ implying $A \cup B \in \cup_{i=1}^{\infty} \mathcal{F}_i$.

(b) Put $\Omega_i = \{j\}_{j=1}^{i}$, and let $\mathcal{F}_i$ be the σ-algebra of the collection of all subsets of Ω_i for $i \in \mathbb{N}$. Suppose that $\cup_{i=1}^{\infty} \mathcal{F}_i$ is also a σ-algebra. Since, for each i, $\{i\} \in \mathcal{F}_i$ and $\mathcal{F}_i \subseteq \cup_{i=1}^{\infty} \mathcal{F}_i$ we have $\{i\} \in \cup_{i=1}^{\infty} \mathcal{F}_i$. Thus, by our primary assumption, $\mathbb{N} = \cup_{i=1}^{\infty} \{i\} \in \cup_{i=1}^{\infty} \mathcal{F}_i$ and, therefore, $\mathbb{N} \in \mathcal{F}_i$ for some i, which implies $\mathbb{N} \subseteq \Omega_i$, a contradiction. Hence, $\cup_{i=1}^{\infty} \mathcal{F}_i$ is not a σ-algebra.□

Exercise 2.7.6. Suppose that $\Omega = [0, 1]$ is the unit interval, and $\mathcal{F}$ is the set of all subsets A such that either A or A^c is finite, and P is defined by $P(A) = 0$ if A is finite, and $P(A) = 1$ if A^c is finite.
(a) Is $\mathcal{F}$ an algebra?
(b) Is $\mathcal{F}$ a σ-algebra?
(c) Is P finitely additive?
(d) Is P countably additive on $\mathcal{F}$ (as the previous exercise)?

Solution. (a) Yes. First, since ϕ is finite and $\Omega^c = \phi$ is finite, we have $\phi, \Omega \in \mathcal{F}$. Second, let $A \in \mathcal{F}$, then either A or A^c is finite implying either A^c or A is finite, hence, $A^c \in \mathcal{F}$. Third, let $A, B \in \mathcal{F}$. Then, we have several cases:
(i) A finite (A^c infinite):
(i-i) B finite (B^c infinite): $A \cup B$ finite , $(A \cup B)^c = A^c \cap B^c$ infinite
(i-ii) B^c finite (B infinite): $A \cup B$ infinite , $(A \cup B)^c = A^c \cap B^c$ finite
(ii) A^c finite (A infinite):
(ii-i) B finite (B^c infinite): $A \cup B$ infinite , $(A \cup B)^c = A^c \cap B^c$ finite
(ii-ii) B^c finite (B infinite): $A \cup B$ infinite , $(A \cup B)^c = A^c \cap B^c$ finite.
Hence, $A \cup B \in \mathcal{F}$.

(b) No. For any $n \in \mathbb{N}$, $\{\frac{1}{n}\} \in \mathcal{F}$. But, $\{\frac{1}{n}\}_{n=1}^{\infty} \notin \mathcal{F}$.

(c) Yes. let $A, B \in \mathcal{F}$ be disjoint. Then, we have several cases:
(i) A finite (A^c infinite):
(i-i) B finite (B^c infinite): $P(A \cup B) = 0 = 0 + 0 = P(A) + P(B)$
(i-ii) B^c finite (B infinite): $P(A \cup B) = 1 = 0 + 1 = P(A) + P(B)$
(ii) A^c finite (A infinite):
(ii-i) B finite (B^c infinite): $P(A \cup B) = 1 = 1 + 0 = P(A) + P(B)$
(ii-ii) B^c finite (B infinite): Since $A^c \cup B^c$ is finite, $A^c \cup B^c \neq \Omega$ implying $A \cap B \neq \phi$.
Hence, $P(A \cup B) = P(A) + P(B)$.

(d) Yes. Let $A_1, A_2, \cdots \in \mathcal{F}$ be disjoint such that $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$. Then, there are two cases:
(i) $\cup_{n=1}^{\infty} A_n$ finite:
In this case, for each $n \in \mathbb{N}$, A_n is finite. Therefore, $P(\cup_{n=1}^{\infty} A_n) = 0 = \sum_{n=1}^{\infty} 0 = \sum_{n=1}^{\infty} P(A_n)$.
(ii) $(\cup_{n=1}^{\infty} A_n)^c$ finite ($\cup_{n=1}^{\infty} A_n$ infinite):
In this case, there is some $n_0 \in \mathbb{N}$ such that A_{n_0} is infinite.

(In fact, if all A_n's are finite, then $\cup_{n=1}^{\infty} A_n$ will be countable. Hence it has Lebesgue measure zero implying that its complement has Lebesgue measure one. On the other hand, its complement is finite having Lebesgue measure zero, a contradiction.)

Now, let $n \neq n_0$, then $A_n \cap A_{n_0} = \phi$ yields $A_n \subseteq A_{n_0}^c$. But $A_{n_0}^c$ is finite, implying that A_n is finite. Therefore:

$$P(\cup_{n=1}^{\infty} A_n) = 1 = 1 + 0 = P(A_{n_0}) + \sum_{n \neq n_0} 0 = P(A_{n_0}) + \sum_{n \neq n_0} P(A_{n_0}) = \sum_{n=1}^{\infty} P(A_n).$$

Accordingly, P is countably additive on $\mathcal{F}$.$\square$

Exercise 2.7.8. For the example of Exercise 2.7.7, is P uncountably additive (cf. page 2)?

Solution. No, if it is uncountably additive, then:

$$1 = P([0,1]) = P(\cup_{x\in[0,1]}\{x\}) = \sum_{x\in[0,1]} P(\{x\}) = \sum_{x\in[0,1]} 0 = 0,$$

a contradiction.□

Exercise 2.7.10. Prove that the collection $\mathcal{J}$ of (2.5.10) is a semi-algebra.

Solution. First, by definition $\phi, \mathbb{R} \in \mathcal{J}$. Second, let $A_1, A_2 \in \mathcal{J}$. If $A_i = \phi(\Omega)$, then $A_1 \cap A_2 = \phi(A_j) \in \mathcal{J}$. Assume, $A_1, A_2 \neq \phi, \Omega$, then we have the following cases:
(i) $A_1 = (-\infty, x_1]$:
(i-i) $A_2 = (-\infty, x_2]$: $A_1 \cap A_2 = (-\infty, \min(x_1, x_2)] \in \mathcal{J}$
(i-ii) $A_2 = (y_2, \infty)$: $A_1 \cap A_2 = (y_2, x_1] \in \mathcal{J}$
(i-iii) $A_2 = (x_2, y_2]$: $A_1 \cap A_2 = (x_2, \min(x_1, y_2)] \in \mathcal{J}$
(ii) $A_1 = (y_1, \infty)$:
(ii-i) $A_2 = (-\infty, x_2]$: $A_1 \cap A_2 = (y_1, x_2] \in \mathcal{J}$
(ii-ii) $A_2 = (y_2, \infty)$: $A_1 \cap A_2 = (\max(y_1, y_2), \infty) \in \mathcal{J}$
(ii-iii) $A_2 = (x_2, y_2]$: $A_1 \cap A_2 = (\max(x_2, y_1), y_2] \in \mathcal{J}$
(iii) $A_1 = (x_1, y_1]$:
(iii-i) $A_2 = (-\infty, x_2]$: $A_1 \cap A_2 = (x_1, \min(x_2, y_1)] \in \mathcal{J}$
(iii-ii) $A_2 = (y_2, \infty)$: $A_1 \cap A_2 = (\max(y_2, x_1), y_1] \in \mathcal{J}$
(iii-iii) $A_2 = (x_2, y_2]$: $A_1 \cap A_2 = (\max(x_1, x_2), \min(y_1, y_2)] \in \mathcal{J}$.
Accordingly, $A_1 \cap A_2 \in \mathcal{J}$. Now, the general case is easily proved by induction (Check!).

Third, let $A \in \mathcal{J}$. If $A = \phi(\Omega)$, then $A^c = \Omega(\phi) \in \mathcal{J}$. If $A = (-\infty, x]$, then $A^c = (x, \infty) \in \mathcal{J}$. If $A = (y, \infty)$, then $A^c = (-\infty, y] \in \mathcal{J}$. Finally, if $A = (x, y]$, then $A^c = (-\infty, x] \cup (y, \infty)$ where both disjoint components are in $\mathcal{J}$.□

Exercise 2.7.12. Let K be the Cantor set as defined in Subsection 2.4. Let $D_n = K \oplus \frac{1}{n}$ where $K \oplus \frac{1}{n}$ is defined as in (1.2.4). Let $B = \cup_{n=1}^{\infty} D_n$.
(a) Draw a rough sketch of D_3.
(b) What is $\lambda(D_3)$?
(c) Draw a rough sketch of B.
(d) What is $\lambda(B)$?

Solution.(a)

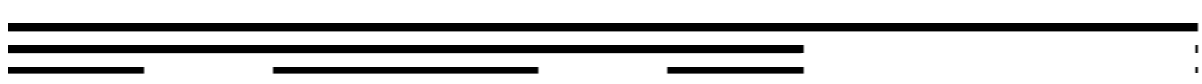


Figure 1: Constructing the sketch of the set $D_3 = K \oplus \frac{1}{3}$

(b) $\lambda(D_3) = \lambda(k \oplus \frac{1}{3}) = \lambda(K) = 0$.

(c)

Figure 2: Constructing the sketch of the set $B = \cup_{n=1}^{\infty} D_n$

In Figure 2, the line one illustrates a rough sketch of the set D_1, the line two illustrates a rough sketch of $\cup_{n=1}^{2} D_n$, the line three illustrates a rough sketch of $\cup_{n=1}^{3} D_n$, and so on.

(d) From $\lambda(D_n) = \lambda(k \oplus \frac{1}{n}) = \lambda(K) = 0$ for all $n \in \mathbb{N}$,and $\lambda(B) \leq \sum_{n=1}^{\infty} \lambda(D_n)$ it follows that $\lambda(B) = 0.\square$

Exercise. 2.7.14. Let $\Omega = \{1, 2, 3, 4\}$, with $\mathcal{F}$ the collection of all subsets of Ω. Let P and Q be two probability measures on$\mathcal{F}$, such that $P\{1\} = P\{2\} = P\{3\} = P\{4\} = \frac{1}{4}$, and $Q\{2\} = Q\{4\} = \frac{1}{2}$, extended to $\mathcal{F}$ by linearity. Finally, let $\mathcal{J} = \{\phi, \Omega, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}\}$.
(a) Prove that $P(A) = Q(A)$ for all $A \in \mathcal{J}$.
(b) Prove that there is $A \in \sigma(\mathcal{J})$ with $P(A) \neq Q(A)$.
(c) Why does this not contradict Proposition 2.5.8?

Solution. (a)

$$P(\phi) = 0 = Q(\phi),$$
$$P(\Omega) = 1 = Q(\Omega),$$
$$P\{a, b\} = P\{a\} + P\{b\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = Q\{a\} + Q\{b\} = Q\{a, b\} \text{ for all } \{a, b\} \in \mathcal{J}.$$

(b) Take $A = \{1, 2, 3\} = \{1, 2\} \cup \{2, 3\} \in \sigma(\mathcal{J})$. Then:

$$P(A) = \sum_{i=1}^{3} P(\{i\}) = \frac{3}{4} \neq \frac{1}{2} = \sum_{i=1}^{3} Q(\{i\}) = Q(A).$$

(c) Since $\{1, 2\}, \{2, 3\} \in \mathcal{J}$ and $\{1, 2\} \cap \{2, 3\} = \{2\} \notin \mathcal{J}$, the set $\mathcal{J}$ is not a semi-algebra. Thus, the hypothesis of the proposition 2.5.8. is not satisfied by $\mathcal{J}.\square$

Exercise 2.7.16. (a) Where in the proof of Theorem 2.3.1. was assumption (2.3.3) used?
(b) How would the conclusion of Theorem 2.3.1 be modified if assumption (2.3.3) were dropped(but all other assumptions remained the same)?

Solution.(a) It was used in the proof of Lemma 2.3.5.

(b) In the assertion of the Theorem 2.3.1, the equality $P^*(A) = P(A)$ will be replaced by $P^*(A) \leq P(A)$ for all $A \in \mathcal{J}.\square$

Exercise 2.7.18. Let $\Omega = \{1, 2\}$, $\mathcal{J} = \{\phi, \Omega, \{1\}\}, P(\phi) = 0, P(\Omega) = 1$, and $P\{1\} = \frac{1}{3}$.
(a) Can Theorem 2.3.1, Corollary 2.5.1, or Corollary 2.5.4 be applied in this case? Why or why not?
(b) Can this P be extended to a valid probability measure? Explain.

Solution.(a) No. Because $\mathcal{J}$ is not a semi-algebra (in fact $\{1\} \in \mathcal{J}$ but $\{1\}^c = \{2\}$ cannot be written as a union of disjoint elements of $\mathcal{J}$.

(b) Yes. It is sufficient to put $\mathcal{M} = \{\phi, \Omega, \{1\}, \{2\}\}$ and $P^\dagger(A) = P(A)$ if $A \in \mathcal{J}, \frac{2}{3}$ if $A = \{2\}$.□

Exercise 2.7.20. Let P and Q be two probability measures defined on the same sample space Ω and σ-algebra $\mathcal{F}$.
(a) Suppose that $P(A) = Q(A)$ for all $A \in \mathcal{F}$ with $P(A) \leq \frac{1}{2}$. Prove that $P = Q$. i.e. that $P(A) = Q(A)$ for all $A \in \mathcal{F}$.
(b) Give an example where $P(A) = Q(A)$ for all $A \in \mathcal{F}$ with $P(A) < \frac{1}{2}$, but such that $P \neq Q$. i.e. that $P(A) \neq Q(A)$ for some $A \in \mathcal{F}$.

Solution.(a) Let $A \in \mathcal{F}$. If $P(A) \leq \frac{1}{2}$, then $P(A) = Q(A)$, by assumption. If $P(A) > \frac{1}{2}$, then $P(A^c) < \frac{1}{2}$. Therefore, $1 - P(A) = P(A^c) = Q(A^c) = 1 - Q(A)$ implying $P(A) = Q(A)$.

(b) Take $\Omega = \{1, 2\}$ and $\mathcal{F} = \{\phi, \{1\}, \{2\}, \{1, 2\}\}$. Define P, Q respectively as follows:

$$P(\phi) = 0, P\{1\} = \frac{1}{2}, P\{2\} = \frac{1}{2}, \text{and } \ P(\Omega) = 1.$$

$$Q(\phi) = 0, Q\{1\} = \frac{1}{3}, Q\{2\} = \frac{2}{3}, \text{and } \ P(\Omega) = 1.$$

□

Exercise 2.7.22. Let $(\Omega_1, \mathcal{F}_1, P_1)$ be Lebesgue measure on $[0, 1]$. Consider a second probability triple $(\Omega_2, \mathcal{F}_2, P_2)$, defined as follows: $\Omega_2 = \{1, 2\}$, $\mathcal{F}_2$ consists of all subsets of Ω_2, and P_2 is defined by $P_2\{1\} = \frac{1}{3}, P_2\{2\} = \frac{2}{3}$, and additivity. Let $(\Omega, \mathcal{F}, P)$ be the product measure of $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$.
(a) Express each of $\Omega, \mathcal{F}$, and P as explicitly as possible.
(b) Find a set $A \in \mathcal{F}$ such that $P(A) = \frac{3}{4}$.

Solution.(a) Define Ω by:

$$\Omega = \Omega_1 \times \Omega_2 = [0, 1] \times \{1, 2\} = \{(w_1, w_2) : 0 \leq w_1 \leq 1, w_2 = 1, 2\}.$$

Now, for $\mathcal{J} = \{A \times \phi, A \times \{1\}, A \times \{2\}, A \times \{1,2\} : A \in \mathcal{F}_1\}$ define the algebra $\mathcal{F} = sigma-algebra(\mathcal{J})$ where:

$$\begin{aligned}
(A \times B)^c &= (A^c \times B^c) \cup (A^c \times B) \cup (A \times B^c) : \\
&\qquad\qquad\qquad\qquad A \in \mathcal{F}, B \subseteq \Omega_2, \\
(A_1 \times B_1) \cap (A_2 \times B_2) &= (A_1 \cap A_2) \times (B_1 \cap B_2) : \\
&\qquad\qquad\qquad\qquad A_1, A_2 \in \mathcal{F}, B_1, B_2 \subseteq \Omega_2, \\
(A_1 \times B_1) \cup (A_2 \times B_2) &= ((A_1 - A_2) \times B_1) \cup ((A_1 \cap A_2) \times (B_1 \cup B_2) \cup ((A_2 - A_1) \times B_2) : \\
&\qquad\qquad\qquad\qquad A_1, A_2 \in \mathcal{F}, B_1, B_2 \subseteq \Omega_2.
\end{aligned}$$

Then:

$$P(A \times B) = 0 \text{ if } B = \phi, \frac{\lambda(A)}{3} \text{ if } B = \{1\}, \frac{2\lambda(A)}{3} \text{ if } B = \{2\}, \text{ and } \lambda(A) \text{ if } B = \{1,2\}.$$

(b) Take $A = [0, \frac{3}{4}] \times \{1, 2\}$, then $P(A) = \lambda[0, \frac{3}{4}] = 3/4.\square$

Chapter 3

Further probabilistic foundations

Exercise 3.6.2. Let $(\Omega, \mathcal{F}, P)$ be Lebesgue measure on $[0,1]$. Let $A = (\frac{1}{2}, \frac{3}{4})$ and $B = (0, \frac{2}{3})$. Are A and B independent events?

Solution. Yes. In this case, $P(A) = \frac{1}{4}$, $P(B) = \frac{2}{3}$, and $P(A \cap B) = P((\frac{1}{2}, \frac{2}{3})) = \frac{1}{6}$. Hence:

$$P(A \cap B) = \frac{1}{6} = \frac{1}{4}\frac{2}{3} = P(A)P(B).$$

□

Exercise 3.6.4. Suppose $\{A_n\} \nearrow A$. Let $f : \Omega \to \mathbb{R}$ be any function. Prove that $\lim_{n\to\infty} \inf_{w \in A_n} f(w) = \inf_{w \in A} f(w)$.

Solution. Given $\epsilon > 0$. Using the definition of infimum, there exists $w_\epsilon \in A$ such that $f(w_\epsilon) \leq \inf_{w \in A} f(w) + \epsilon$. On the other hand, $A = \cup_{n=1}^{\infty} A_n$ and $A_n \nearrow A$, therefor, there exists $N \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n \geq N$ we have $w_\epsilon \in A_n$, implying $\inf_{w \in A_n} f(w) \leq f(w_\epsilon)$. Combining two recent results yields:

$$\inf_{w \in A_n} f(w) < \inf_{w \in A} f(w) + \epsilon \qquad n = N, N+1, ...(\star)$$

Next, since $A_N \subseteq A_{N+1} \subseteq \cdots \subseteq A$, for any $n \in \mathbb{N}$ with $n \geq N$ we have $\inf_{w \in A} f(w) - \epsilon < \inf_{w \in A} f(w) \leq \inf_{w \in A_n} f(w)$. Accordingly:

$$\inf_{w \in A} f(w) - \epsilon < \inf_{w \in A_n} f(w) \qquad n = N, N+1, ...(\star\star)$$

Finally, by $(\star)$ and $(\star\star)$:

$$|\inf_{w \in A_n} f(w) - \inf_{w \in A} f(w)| < \epsilon \qquad n = N, N+1, ...,$$

proving the assertion.□

Exercise 3.6.6. Let X, Y, and Z be three independent random variables, and set $W = X + Y$. Let

$B_{k,n} = \{(n-1)2^{-k} \leq X < n2^{-k}\}$ and let $C_{k,m} = \{(m-1)2^{-k} \leq Y < m2^{-k}\}$. Let

$$A_k = \bigcup_{n,m\in\mathbb{Z}:(n+m)2^{-k}<x} (B_{k,n} \cap C_{k,m}).$$

Fix $x, z \in \mathbb{R}$, and let $A = \{X + Y < x\} = \{W < x\}$ and $D = \{Z < z\}$.
(a) Prove that $\{A_k\} \nearrow A$.
(b) Prove that A_k and D are independent.
(c) By continuity of probabilities, prove that A and D are independent.
(d) Use this to prove that W and Z are independent.

Solution.(a) First, let $k \geq 0$ be fixed, and $I_{k,n} = [\frac{n-1}{2^k}, \frac{n}{2^k})$ and $J_{k,m} = [\frac{m-1}{2^k}, \frac{m}{2^k})$. Then, $B_{k,n} = X^{-1}(I_{k,n})$ and $C_{k,m} = Y^{-1}(J_{k,m})$. Accordingly, by increasing the number of events under the unions it follows that:

$$\begin{aligned}
A_k &= \bigcup_{n,m\in\mathbb{Z}:\frac{n+m}{2^k}<x} (B_{k,n} \cap C_{k,m}) = \bigcup_{n,m\in\mathbb{Z}:\frac{n+m}{2^k}<x} (X^{-1}(I_{k,n}) \cap Y^{-1}(J_{k,m})) \\
&= \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n-1} \cup I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m-1} \cup J_{k+1,2m})) \\
&= \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} ((X^{-1}(I_{k+1,2n-1}) \cup X^{-1}(I_{k+1,2n})) \cap (Y^{-1}(J_{k+1,2m-1}) \cup (Y^{-1}(J_{k+1,2m})) \\
&= \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} \Big((X^{-1}(I_{k+1,2n-1}) \cap Y^{-1}(J_{k+1,2m-1}))) \cup (X^{-1}(I_{k+1,2n-1}) \cap Y^{-1}(J_{k+1,2m})) \\
&\qquad \cup (X^{-1}(I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m-1})) \cup (X^{-1}(I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m}))\Big) \\
&= \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n-1}) \cap Y^{-1}(J_{k+1,2m-1})) \cup \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n-1}) \cap Y^{-1}(J_{k+1,2m})) \\
&\quad \cup \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m-1})) \cup \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m})) \\
&\subseteq \bigcup_{n,m\in\mathbb{Z}:\frac{2n-1+2m-1}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n-1}) \cap Y^{-1}(J_{k+1,2m-1})) \cup \bigcup_{n,m\in\mathbb{Z}:\frac{2n-1+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n-1}) \cap Y^{-1}(J_{k+1,2m})) \\
&\quad \cup \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m-1}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m-1})) \cup \bigcup_{n,m\in\mathbb{Z}:\frac{2n+2m}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,2n}) \cap Y^{-1}(J_{k+1,2m})) \\
&= \bigcup_{n^*,m^*\in\mathbb{Z}:\frac{n^*+m^*}{2^{(k+1)}}<x} (X^{-1}(I_{k+1,n^*}) \cap Y^{-1}(J_{k+1,m^*})) \\
&= \bigcup_{n^*,m^*\in\mathbb{Z}:\frac{n^*+m^*}{2^{(k+1)}}<x} (B_{k+1,n^*} \cap C_{k+1,m^*}) \\
&= A_{k+1},
\end{aligned}$$

implying (by induction):

$$A_0 \subseteq A_1 \subseteq \cdots \subseteq A_k \subseteq \cdots .$$

Second, define $K_{k,p} = [\frac{p-1}{2^k}, \frac{p}{2^k})$, and $D_{k,p} = (X+Y)^{-1}(K_{k,p})$. Then, by uniqueness of m, n, p in the related partitions it follows that:

$$\begin{aligned}
\bigcup_{k=0}^{\infty} A_k &= \bigcup_{k=0}^{\infty}(\bigcup_{n,m\in\mathbb{Z}: \frac{n+m}{2^k}<x} (B_{k,n}\cap C_{k,m})) \\
&= \bigcup_{k=0}^{\infty}(\bigcup_{n,m\in\mathbb{Z}: \frac{n+m}{2^k}<x} (X^{-1}(I_{k,n})\cap Y^{-1}(J_{k,m}))) \\
&= \bigcup_{k=0}^{\infty}(\bigcup_{p\in\mathbb{Z}: \frac{p}{2^k}<x} \bigcup_{n\in\mathbb{Z}} \bigcup_{m\in\mathbb{Z}} ((X+Y)^{-1}(K_{k,p})\cap X^{-1}(I_{k,n})\cap Y^{-1}(J_{k,m}))) \\
&= (\bigcup_{k=0}^{\infty}(\bigcup_{p\in\mathbb{Z}: \frac{p}{2^k}<x} (X+Y)^{-1}(K_{k,p})) \bigcap (\bigcup_{k=0}^{\infty} \bigcup_{n\in\mathbb{Z}} X^{-1}(I_{k,n})) \bigcap (\bigcup_{k=0}^{\infty} \bigcup_{m\in\mathbb{Z}} Y^{-1}(J_{k,m})) \\
&= \{X+Y<x\}\cap\{-\infty<X<\infty\}\cap\{-\infty<Y<\infty\} \\
&= A.
\end{aligned}$$

Accordingly, by above two results the plausible statement follows.

(b)For any $k\in\mathbb{N}$ we have that:

$$\begin{aligned}
P(A_k\cap D) &= P(\bigcup_{n,m\in\mathbb{Z}: \frac{(n+m)}{2^k}<x} (B_{k,n}\cap C_{k,m}\cap D)) \\
&= \sum_{n,m\in\mathbb{Z}: \frac{(n+m)}{2^k}<x} P(B_{k,n}\cap C_{k,m}\cap D) \\
&= \sum_{n,m\in\mathbb{Z}: \frac{(n+m)}{2^k}<x} P(B_{k,n}\cap C_{k,m})P(D) \\
&= P(A_k)P(D).
\end{aligned}$$

(c) Using part (b):

$$P(A\cap D) = \lim_k P(A_k\cap D) = \lim_k (P(A_k)P(D)) = P(A)P(D).$$

(d) This is the consequence of part (c) and Proposition 3.2.4.□

Exercise 3.6.8. Let λ be Lebesgue measure on $[0,1]$, and let $0\le a\le b\le c\le d\le 1$ be arbitrary real numbers with $d\ge b+c-a$. Give an example of a sequence $A_1, A_2, \cdots$ of intervals in $[0,1]$, such that $\lambda(\liminf_n A_n) = a, \liminf_n \lambda(A_n) = b, \limsup_n \lambda(A_n) = c$, and $\lambda(\limsup_n A_n) = d$. For bonus points, solve the question when $d < b+c-a$, with each A_n a finite union of intervals.

Solution. Let $e = (d+a)-(b+c)$, and consider:

$$\begin{aligned}
A_{3n} &= (0, b+e), \\
A_{3n-1} &= (e, b+e), \\
A_{3n-2} &= (b-a+e, c+b-a+e),
\end{aligned}$$

for all $n \in \mathbb{N}$. Then:

$$\begin{aligned}
\lambda(\liminf_n A_n) &= \lambda(b-a+e, b+e) = a,\\
\liminf_n \lambda(A_n) &= \lambda(e, b+e) = b,\\
\limsup_n \lambda(A_n) &= \lambda(b-a+e, c+b-a+e) = c,\\
\lambda(\limsup_n A_n) &= \lambda(0, d) = d,
\end{aligned}$$

where $b+e \leq c$ or $d \leq 2c-a$.$\Box$

Exercise 3.6.10. Let $A_1, A_2, \cdots$ be a sequence of events, and let $N \in \mathbb{N}$. Suppose there are events B and C such that $B \subseteq A_n \subseteq C$ for all $n \geq N$, and such that $P(B) = P(C)$. Prove that $P(\liminf_n A_n) = P(\limsup_n A_n) = P(B) = P(C)$.

Solution. Since:

$$B \subseteq \cap_{n=N}^{\infty} A_n \subseteq \cup_{m=1}^{\infty} \cap_{n=m}^{\infty} A_n \subseteq \cap_{m=1}^{\infty} \cup_{n=m}^{\infty} A_n \subseteq \cup_{n=N}^{\infty} A_n \subseteq C,$$

$P(B) \leq P(\liminf_n A_n) \leq P(\limsup_n A_n) \leq P(C)$. Now, using the condition $P(B) = P(C)$, yields the desired result.$\Box$

Exercise 3.6.12. Let X be a random variable with $P(X > 0) > 0$. Prove that there is a $\delta > 0$ such that $P(X \geq \delta) > 0$.[Hint: Don't forget continuity of probabilities.]

Solution. Method (1):
Put $A = \{X > 0\}$ and $A_n = \{X \geq \frac{1}{n}\}$ for all $n \in \mathbb{N}$. Then, $A_n \nearrow A$ and using proposition 3.3.1, $\lim_n P(A_n) = P(A)$. But, $P(A) > 0$, therefore , there is $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geq N$ we have $P(A_n) > 0$. In particular, $P(A_N) > 0$. Take, $\delta = \frac{1}{N}$.
Method (2):
Put $A = \{X > 0\}$ and $A_n = \{X \geq \frac{1}{n}\}$ for all $n \in \mathbb{N}$. Then, $A = \cup_{n=1}^{\infty} A_n$ and $P(A) \leq \sum_{n=1}^{\infty} P(A_n)$. If for any $n \in \mathbb{N}$, $P(A_n) = 0$, then using recent result, $P(A) = 0$, a contradiction. Therefore, there is at least one $N \in \mathbb{N}$ such that $P(A_N) > 0$. Take, $\delta = \frac{1}{N}$.$\Box$

Exercise 3.6.14. Let $\delta, \epsilon > 0$, and let $X_1, X_2, \cdots$ be a sequence of independent non-negative random variables such that $P(X_i \geq \delta) \geq \epsilon$ for all i. Prove that with probability one, $\sum_{i=1}^{\infty} X_i = \infty$.

Solution. Since $P(X_i \geq \delta) \geq \epsilon$ for all i, $\sum_{i=1}^{\infty} P(X_i \geq \delta) = \infty$, and by Borel-Cantelli Lemma:

$$P(\limsup_i (X_i \geq \delta)) = 1.(\star)$$

On the other hand,

$$\limsup_i (X_i \geq \delta) \subseteq (\sum_{i=1}^{\infty} X_i = \infty)$$

(in fact, let $w \in \limsup_i (X_i \geq \delta)$, then there exists a sequence $\{i_j\}_{j=1}^{\infty}$ such that $X_{i_j} \geq \delta$ for all $j \in \mathbb{N}$, yielding $\sum_{j=1}^{\infty} X_{i_j}(w) = \infty$, and consequently, $\sum_{i=1}^{\infty} X_i(w) = \infty$. This implies $w \in (\sum_{i=1}^{\infty} X_i = \infty)$).

Consequently:

$$P(\limsup_i (X_i \geq \delta)) \leq P(\sum_{i=1}^{\infty} X_i = \infty).(\star\star)$$

Now, by $(\star)$ and $(\star\star)$ it follows $P(\sum_{i=1}^{\infty} X_i = \infty) = 1$. □

Exercise 3.6.16. Consider infinite, independent, fair coin tossing as in subsection 2.6, and let H_n be the event that the n^{th} coin is heads. Determine the following probabilities.
(a) $P(\cap_{i=1}^{9} H_{n+i}\ i.o.)$.
(b) $P(\cap_{i=1}^{n} H_{n+i}\ i.o.)$.
(c) $P(\cap_{i=1}^{[2\log_2 n]} H_{n+i}\ i.o.)$.
(d)Prove that $P(\cap_{i=1}^{[\log_2 n]} H_{n+i}\ i.o.)$ must equal either 0 or1.
(e) Determine $P(\cap_{i=1}^{[\log_2 n]} H_{n+i}\ i.o.)$.[Hint: Find the right subsequence of indices.]

Solution.(a) First of all, put $A_n = \cap_{i=1}^{9} H_{n+i}, (n \geq 1)$, then :

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} (\frac{1}{2})^9 = \infty.$$

But the events $A_n, (n \geq 1)$ are not independent. Consider the independent subsequence $B_n = A_{f(n)}, (n \geq 1)$ where the function f is given by $f(n) = 10n, (n \geq 1)$. Besides,

$$\sum_{n=1}^{\infty} P(B_n) = \sum_{n=1}^{\infty} P(\cap_{i=1}^{9} H_{10n+i}) = \sum_{n=1}^{\infty} (\frac{1}{2})^9 = \infty.$$

Now,using Borel-Cantelli Lemma, $P(B_n\ i.o.) = 1$ implying $P(A_n\ i.o.) = 1$.

(b) Put $A_n = \cap_{i=1}^{n} H_{n+i}, (n \geq 1)$. Since

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} (\frac{1}{2})^n = 1 < \infty,$$

the Borel-Cantelli Lemma implies $P(A_n i.o.) = 0$.
(c) Put $A_n = \cap_{i=1}^{[2\log_2 n]} H_{n+i}, (n \geq 1)$. Since

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} (\frac{1}{2})^{[2\log_2 n]} \leq \sum_{n=1}^{\infty} (\frac{1}{n})^2 < \infty,$$

Borel-Cantelli Lemma implies $P(A_n\ i.o.) = 0$.

(d),(e) Put $A_n = \cap_{i=1}^{[\log_2 n]} H_{n+i}, (n \geq 1)$, then :

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} (\frac{1}{2})^{[\log_2 n]} = \infty.$$

But the events $A_n, (n \geq 1)$ are not independent. Consider the independent subsequence $B_n = A_{f(n)}, (n \geq 1)$ where the function f is given by $f(n) = [n\log_2(n^2)], (n \geq 1)$. In addition,

$$\sum_{n=1}^{\infty} P(B_n) = \sum_{n=1}^{\infty} P(\cap_{i=1}^{[\log_2 f(n)]} H_{f(n)+i}) = \sum_{n=1}^{\infty} (\frac{1}{2})^{\log_2 f(n)} = \infty.$$

Using Borel-Cantelli Lemma, $P(B_n \ i.o.) = 1$ implying $P(A_n \ i.o.) = 1$.□

Exercise. 3.6.18. Let $A_1, A_2, \cdots$ be any independent sequence of events, and let $S_x = \{\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{A_i} \leq x\}$. Prove that for each $x \in \mathbb{R}$ we have $P(S_x) = 0$ or 1.

Solution. For a fixed $(m \geq 1)$, we have:

$$S_x = \left\{\liminf_{n\to\infty} \frac{1}{n}\sum_{i=m}^{n} 1_{A_i} \leq x\right\} = \bigcap_{s=1}^{\infty}\bigcup_{N=m}^{\infty}\bigcap_{n=N}^{\infty}\left\{\frac{1}{n}\sum_{i=m}^{n} 1_{A_i} \leq x + \frac{1}{s}\right\},$$

and

$$\left\{\frac{1}{n}\sum_{i=m}^{n} 1_{A_i} \leq x + \frac{1}{s}\right\} \in \sigma(A_m, A_{m+1}, \cdots),$$

which imply that

$$S_x \in \sigma(A_m, A_{m+1}, \cdots),$$

yielding: $S_x \in \bigcap_{m=1}^{\infty} \sigma(A_m, A_{m+1}, \cdots)$. Consequently, by Theorem (3.5.1), $P(S_x) = 0$ or 1.□
In this chapter, we only define the tail field over a sequence of events, instead of the random variables 1_{A_m}...

Chapter 4

Expected values

Exercise 4.5.2. Let X be a random variable with finite mean, and let $a \in \mathbb{R}$ be any real number. Prove that $E(\max(X, a)) \geq \max(E(X), a)$.

Solution. Since $E(.)$ is order preserving, from

$$\max(X, a) \geq X$$

it follows $E(\max(X, a)) \geq E(X)$. Similarly, from

$$\max(X, a) \geq a$$

it follows that $E(\max(X, a)) \geq E(a) = a$.
Combining the recent result yields,

$$E(\max(X, a)) \geq \max(E(X), a).$$

□

Exercise 4.5.4. Let $(\Omega, \mathcal{F}, P)$ be the uniform distribution on $\Omega = \{1, 2, 3\}$, as in Example 2.2.2. Find random variables X, Y, and Z on $(\Omega, \mathcal{F}, P)$ such that $P(X > Y)P(Y > Z)P(Z > X) > 0$, and $E(X) = E(Y) = E(Z)$.

Solution. Put:

$$X = 1_{\{1\}}, Y = 1_{\{2\}}, Z = 1_{\{3\}}.$$

Then,

$$E(X) = E(Y) = E(Z) = \frac{1}{3}.$$

Besides, $P(X > Y) = P(Y > Z) = P(Z > X) = \frac{1}{3}$, implying:

$$P(X > Y)P(Y > Z)P(Z > X) = (\frac{1}{3})^3 > 0.$$

□

Exercise 4.5.6. Let X be a random variable defined on Lebesgue measure on $[0, 1]$, and suppose that X is a one to one function, i.e. that if $w_1 = w_2$ then $X(w_1) \neq X(w_2)$. Prove that X is not a simple random variable.

Solution. Suppose X be a simple random variable. Since $|X([0, 1])| < \aleph_0 < c = |[0, 1]|$ (where $||$ refers to the cardinality of the considered sets), we conclude that there is at least one $y \in X([0,1])$ such that for at least two elements $w_1, w_2 \in [0,1]$ we have $X(w_1) = y = X(w_2)$, contradicting injectivity of X.□

Exercise 4.5.8. Let $f(x) = ax^2 + bx + c$ be a second degree polynomial function (where $a, b, c \in \mathbb{R}$ are constants).
(a) Find necessary and sufficient conditions on a, b, and c such that the equation $E(f(\alpha X)) = \alpha^2 E(f(X))$ holds for all $\alpha \in \mathbb{R}$ and all random variables X.
(b) Find necessary and sufficient conditions on a, b, and c such that the equation $E(f(X - \beta)) = E(f(X))$ holds for all $\beta \in \mathbb{R}$ and all random variables X.
(c) Do parts (a) and (b) account for the properties of the variance function? Why or why not?

Solution. (a) Let for $f(x) = ax^2 + bx + c$ we have $E(f(\alpha X)) = \alpha^2 E(f(X))$ for all $\alpha \in \mathbb{R}$ and all random variables X. Then, a straightforward computation shows that the recent condition is equivalent to :

$$\forall \alpha \forall X : (b\alpha - \alpha^2 b)E(X) + (1 - \alpha^2)E(c) = 0.$$

Consider a random variable X with $E(X) \neq 0$. Put $\alpha = -1$, we obtain $b = 0$. Moreover, put $\alpha = 0$ we obtain $c = 0$. Hence,

$$f(x) = ax^2.$$

Conversely, if $f(x) = ax^2$ then a simple calculation shows that $E(f(\alpha X)) = \alpha^2 E(f(X))$ for all $\alpha \in \mathbb{R}$ and all random variables X (Check!).

(b) Let for $f(x) = ax^2 + bx + c$ we have $E(f(X - \beta)) = E(f(X))$ for all $\beta \in \mathbb{R}$ and all random variables X. Then, a straightforward computation shows that the recent condition is equivalent to :

$$\forall \beta \forall X : (-2a\beta)E(X) + (a\beta^2 - b\beta) = 0.$$

Consider a random variable X with $E(X) = 0$. Then, for any $\beta \in \mathbb{R}$ we have $(a\beta^2 - b\beta) = 0$, implying $a = b = 0$. Hence,

$$f(x) = c.$$

Conversely, if $f(x) = c$ then a simple calculation shows that $E(f(X - \beta)) = E(f(X))$ for all $\beta \in \mathbb{R}$ and all random variables X (Check!).

(c) No. Assume, to reach a contradiction, that $Var(X)$ can be written in the form $E(f(X))$ for some $f(x) = ax^2 + bx + c$. Then:

$$\forall X : E(X^2) - E^2(X) = aE(X^2) + bE(X) + c.(\star)$$

Consider a random variable X with $E(X) = E(X^2) = 0$. Substituting it in $(\star)$ implies $c = 0$. Second, consider a random variable X with $E(X) = 0$ and $E(X^2) \neq 0$. Substituting it in $(\star)$ implies $a = 1$. Now consider two random variables X_1, and X_2 with $E(X_1) = 1$ and $E(X_2) = -1$, respectively. Substituting them in $(\star)$ implies $b = 1$ and $b = -1$, respectively. Thus, 1=-1, a contradiction.□

Exercise 4.5.10. Let $X_1, X_2, \cdots$ be i.i.d. with mean μ and variance σ^2, and let N be an integer-valued random variable with mean m and variance ν, with N independent of all the X_i. Let $S = X_1 + \cdots + X_N = \sum_{i=1}^{\infty} X_i 1_{N \geq i}$. Compute $Var(S)$ in terms of μ, σ^2, m, and ν.

Solution. Method (1):

Need to remove the constant 2 before the summation $\sum_{i \neq j}$

We compute the components of $Var(S) = E(S^2) - E(S)^2$ as follows:

$$\begin{aligned}
E(S^2) &= E((\sum_{i=1}^{\infty} X_i 1_{N\geq i})^2) = E(\sum_{i=1}^{\infty} X_i^2 1_{N\geq i}^2 + \sum_{i\neq j} X_i 1_{N\geq i} X_j 1_{N\geq j}) \\
&= \sum_{i=1}^{\infty} E(X_i^2)E(1_{N\geq i}^2) + \sum_{i\neq j} E(X_i)E(X_j)E(1_{N\geq i}.1_{N\geq j}) \\
&= \sum_{i=1}^{\infty} E(X_i^2)E(1_{N\geq i}^2) + \mu^2.\sum_{i\neq j} E(1_{N\geq i}.1_{N\geq j}), \qquad (*)
\end{aligned}$$

and,

$$\begin{aligned}
E(S)^2 &= (E(\sum_{i=1}^{\infty} X_i 1_{N\geq i}))^2 = (\sum_{i=1}^{\infty} E(X_i)E(1_{N\geq i}))^2 \\
&= \sum_{i=1}^{\infty} E(X_i)^2 E(1_{N\geq i})^2 + \sum_{i\neq j} E(X_i)E(1_{N\geq i})E(X_j)E(1_{N\geq j}) \\
&= \sum_{i=1}^{\infty} E(X_i)^2 E(1_{N\geq i})^2 + \mu^2.\sum_{i\neq j} E(1_{N\geq i})E(1_{N\geq j}). \qquad (**)
\end{aligned}$$

Hence, by $(*)$ and $(**)$ it follows that:

$$\begin{aligned}
Var(S) &= \sum_{i=1}^{\infty} E(X_i^2)E(1_{N\geq i}^2) - \sum_{i=1}^{\infty} E(X_i)^2 E(1_{N\geq i})^2 + \mu^2.\sum_{i\neq j} Cov(1_{N\geq i}, 1_{N\geq j}) \\
&= (\sigma^2 + \mu^2)\sum_{i=1}^{\infty} E(1_{N\geq i}) - \mu^2 \sum_{i=1}^{\infty} E(1_{N\geq i})^2 + \mu^2.\sum_{i\neq j} Cov(1_{N\geq i}, 1_{N\geq j}) \\
&= \sigma^2 \sum_{i=1}^{\infty} E(1_{N\geq i}) + \mu^2 \sum_{i=1}^{\infty} (E(1_{N\geq i}) - E(1_{N\geq i})^2) + \mu^2.\sum_{i\neq j} Cov(1_{N\geq i}, 1_{N\geq j}) \\
&= \sigma^2 \sum_{i=1}^{\infty} E(1_{N\geq i}) + \mu^2 \sum_{i=1}^{\infty} Var(1_{N\geq i}) + \mu^2.\sum_{i\neq j} Cov(1_{N\geq i}, 1_{N\geq j}) \\
&= \sigma^2 E(N) + \mu^2.(Var(\sum_{i=1}^{\infty} 1_{N\geq i}) - \sum_{i\neq j} Cov(1_{N\geq i}, 1_{N\geq j})) + \mu^2.\sum_{i\neq j} Cov(1_{N\geq i}, 1_{N\geq j}) \\
&= \sigma^2 E(N) + \mu^2.Var(N) \\
&= \sigma^2.m + \mu^2.\nu.
\end{aligned}$$

Method(2):

A simple version of theorem 13.3.1 (p.157) is the following equation:

$$Var(Y) = E(Var(Y|X)) + Var(E(Y|X)). \qquad (\dagger)$$

Now, for $Y = S$ and $X = N$ in $(\dagger)$ we have the following arguments:

First, by $E(S|N = n) = E(\sum_{i=1}^{n} X_i) = n.\mu$, we have $E(S|N) = N.\mu$. Consequently:

$$Var(E(S|N)) = Var(N\mu) = \mu^2.Var(N) = \mu^2.\nu. \qquad (\dagger\dagger)$$

Second, by $Var(S|N=n) = Var(\sum_{i=1}^n X_i) = n.\sigma^2$, we have $Var(S|N) = N.\sigma^2$. Hence,

$$E(Var(S|N)) = E(N.\sigma^2) = \sigma^2.m. \qquad (\dagger\dagger\dagger)$$

Accordingly, by $(\dagger), (\dagger\dagger\dagger)$, and $(\dagger\dagger\dagger)$, it follows that:

$$Var(S) = \sigma^2.m + \mu^2.\nu.$$

□

Exercise. 4.5.12. Let X and Y be independent general nonnegative random variables, and let $X_n = \Psi_n(X)$, where $\Psi_n(x) = \min(n, 2^{-n}\lfloor 2^n x\rfloor)$ as in proposition 4.2.5.
(a) Give an example of a sequence of functions $\Phi_n : [0,\infty) \to [0,\infty)$, other that $\Phi_n(x) = \Psi_n(x)$, such that for all x, $0 \leq \Phi_n(x) \leq x$ and $\Phi_n(x) \nearrow x$ as $n \to \infty$.
(b) Suppose $Y_n = \Phi_n(Y)$ with Φ_n as in part (a). Must X_n and Y_n be independent?
(c) Suppose $\{Y_n\}$ is an arbitrary collection of non-negative simple random variables such that $Y_n \nearrow Y$. Must X_n and Y_n be independent?
(d) Under the assumptions of part (c), determine(with proof) which quantities in equation (4.2.7) are necessarily equal.

Solution. (a) Put:

$$\Phi_n(x) = \sum_{m=0}^{\infty}(f_n(x-m)+m)1_{[m,m+1)}(x),$$

where

$$f_n(x) = \sum_{k=0}^{2^{n-1}-1}(2^{n-1}(x-\frac{k}{2^{n-1}})^2 + \frac{k}{2^{n-1}})1_{[\frac{k}{2^{n-1}},\frac{k+1}{2^{n-1}})}(x)$$

for all $0 \leq x < 1$,and $n = 1, 2, \cdots$.
Then, Φ_n has all the required properties (Check!).

(b) Since $\Phi_n(Y)$ is a Borel measurable function, by proposition 3.2.3, $X_n = \Psi_n(X)$ and $Y_n = \Phi_n(Y)$ are independent random variables.

(c) No. It is sufficient to consider:

$$Y_n = \max(\Psi_n(Y) - \frac{1}{n^2}X_n, 0),$$

for all $n \in \mathbb{N}$.

(d) Since $\{X_n\} \nearrow X$, $\{Y_n\} \nearrow Y$ and $\{X_nY_n\} \nearrow XY$, using Theorem 4.2.2., $\lim_n E(X_n) = E(X)$,$\lim_n E(Y_n) = E(Y)$ and $\lim_n E(X_nY_n) = E(XY)$. Hence:

$$\lim_n E(X_n)E(Y_n) = E(X)E(Y)$$

and

$$\lim_n E(X_nY_n) = E(XY).$$

□

Exercise. 4.5.14. Let $Z_1, Z_2, \cdots$ be general random variables with $E(|Z_i|) < \infty$, and let $Z = Z_1 + Z_2 + \cdots$.
(a) Suppose $\sum_i E(Z_i^+) < \infty$ and $\sum_i E(Z_i^-) < \infty$. Prove that $E(Z) = \sum_i E(Z_i)$.
(b) Show that we still have $E(Z) = \sum_i E(Z_i)$ if we have at least one of $\sum_i E(Z_i^+) < \infty$ or $\sum_i E(Z_i^-) < \infty$.
(c) Let $\{Z_i\}$ be independent, with $P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$ for each i. Does $E(Z) = \sum_i E(Z_i)$ in this case? How does that relate to (4.2.8)?

Solution.(a)

$$\begin{aligned}
E(Z) &= E(\sum_i Z_i) \\
&= E(\sum_i (Z_i^+ - Z_i^-)) \\
&= E(\sum_i Z_i^+ - \sum_i Z_i^-) \\
&= E(\sum_i Z_i^+) - E(\sum_i Z_i^-) \\
&= \sum_i E(Z_i^+) - \sum_i E(Z_i^-) \\
&= \sum_i (E(Z_i^+) - E(Z_i^-)) \\
&= \sum_i E(Z_i).
\end{aligned}$$

(b) We prove the assertion for the case $\sum_i E(Z_i^+) = \infty$ and $\sum_i E(Z_i^-) < \infty$ (the proof of other case is analogous.). Similar to part (a) we have:

$$\begin{aligned}
E(Z) &= E(\sum_i Z_i) \\
&= E(\sum_i (Z_i^+ - Z_i^-)) \\
&= E(\sum_i Z_i^+ - \sum_i Z_i^-) \\
&= E(\sum_i Z_i^+) - E(\sum_i Z_i^-) \\
&= \sum_i E(Z_i^+) - \sum_i E(Z_i^-) \\
&= \infty \\
&= \sum_i (E(Z_i^+) - E(Z_i^-)) \\
&= \sum_i E(Z_i).
\end{aligned}$$

(c) Since $E(Z_i) = 0$ for all i, $\sum_i E(Z_i) = 0$. On the other hand, $E(Z)$ is undefined, hence $E(Z) \neq \sum_i E(Z_i)$. This example shows if $\{X_n\}_{n=1}^{\infty}$ are not non-negative, then the equation (4.2.8) may fail.□

Chapter 5

Inequalities and convergence

Exercise. 5.5.2. Give an example of a random variable X and $\alpha > 0$ such that $P(X \geq \alpha) > E(X)/\alpha$.[Hint: Obviously X cannot be non-negative.] Where does the proof of Markov's inequality break down in this case?

Solution. Part one: Let $(\Omega, \mathcal{F}, P)$ be the Lebesgue measure on [0,1]. Define $X : [0,1] \to \mathbb{R}$ by

$$X(w) = (1_{[0,\frac{1}{2}]} - 1_{(\frac{1}{2},1]})(w).$$

Then, by Theorem 4.4, $E(X) = \int_0^1 X(w)dw = 0$. However,

$$P(X \geq \frac{1}{2}) = \frac{1}{2} > 0 = E(X)/\frac{1}{2}.$$

Part two: In the definition of Z we will not have $Z \leq X$.$\square$

Exercise 5.5.4. Suppose X is a nonnegative random variable with $E(X) = \infty$. What does Markov's inequality say in this case?

Solution. In this case, it will be reduced to the trivial inequality $P(X \geq \alpha) \leq \infty$.$\square$

Exercise 5.5.6. For general jointly defined random variables X and Y, prove that $|\mathrm{Corr}(X,Y)| \leq 1$.[Hint: Don't forget the Cauchy-Schwarz inequality.]

Solution. Method(1):
By Cauchy-Schwarz inequality:

$$|\mathrm{Corr}(X,Y)| = |\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}| = |\frac{E((X-\mu_X)(Y-\mu_y))}{\sqrt{E((X-\mu_X)^2)E((Y-\mu_y)^2)}}| \leq \frac{E(|(X-\mu_X)(Y-\mu_y)|)}{\sqrt{E((X-\mu_X)^2)E((Y-\mu_y)^2)}} \leq 1.$$

Method (2):
Since:

$$0 \leq Var(\frac{X}{\sqrt{Var(X)}} + \frac{Y}{\sqrt{Var(Y)}}) = 1 + 1 + 2\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = 2(1 + \mathrm{Corr}(X,Y))$$

we conclude:

$$-1 \leq \mathrm{Corr}(X,Y).(\star)$$

On the other hand, from:

$$0 \leq Var(\frac{X}{\sqrt{Var(X)}} - \frac{Y}{\sqrt{Var(Y)}}) = 1 + 1 - 2\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = 2(1 - \text{Corr}(X,Y)),$$

if follows:

$$\text{Corr}(X,Y) \leq 1.(\star\star)$$

Accordingly, by $(\star)$ and $(\star\star)$ the desired result follows.□

Exercise 5.5.8. Let $\phi(x) = x^2$.
(a) Prove that ϕ is a convex function.
(b) What does Jensen's inequality say for this choice of ϕ?
(c) Where in the text have we already seen the result of part (b)?

Solution.(a) Let ϕ have a second derivative at each point of (a,b). Then ϕ is convex on (a,b) if and only if $\phi''(x) \geq 0$ for each $x \in (a,b)$. Since in this problem $\phi''(x) = 2 \geq 0$, using the recent proposition it follows that ϕ is a convex function.

(b) $E(X^2) \geq E^2(X)$.

(c) We have seen it in page 44, as the first property of $Var(X)$.□

Exercise 5.5.10. Let $X_1, X_2, \cdots$ be a sequence of random variables, with $E(X_n) = 8$ and $Var(X_n) = 1/\sqrt{n}$ for each n. Prove or disprove that $\{X_n\}$ must converge to 8 in probability.

Solution. Given $\epsilon > 0$. Using Proposition 5.1.2:

$$P(|X_n - 8| \geq \epsilon) \leq Var(X_n)/\epsilon^2 = 1/\sqrt{n}\epsilon^2,$$

for all $n \in \mathbb{N}$. Let $n \to \infty$, then:

$$\lim_{n\to\infty} P(|X_n - 8| \geq \epsilon) = 0.$$

Hence, $\{X_n\}$ must converge to 8 in probability.□

Exercise 5.5.12. Give (with proof) an example of two discrete random variables having the same mean and the same variance, but which are not identically distributed.

Solution. As the first example, let

$$\Omega = \{1,2,3,4\}, \mathcal{F} = \mathcal{P}(\Omega), P(X = i) = p_i$$

and $P(Y = i) = q_i$ where

$$(p_1, p_2, p_3, p_4) = (\frac{8}{96}, \frac{54}{96}, \frac{12}{96}, \frac{22}{96})$$

and

$$(q_1, q_2, q_3, q_4) = (\frac{24}{96}, \frac{6}{96}, \frac{60}{96}, \frac{6}{96}).$$

Then, $E(X)=\frac{240}{96}=E(Y)$ and $E(X^2)=\frac{684}{96}=E(Y^2)$ but $E(X^3)=\frac{2172}{96}\neq\frac{2076}{96}=E(Y^3)$.

As the second example, let

$$\Omega_1=\{1,2\}, \mathcal{F}_1=\mathcal{P}(\Omega_1), P(X=1)=\frac{1}{2}=P(X=-1),$$

and

$$\Omega_2=\{-2,0,2\}, \mathcal{F}_2=\mathcal{P}(\Omega_2), P(Y=2)=\frac{1}{8}=P(Y=-2), P(Y=0)=\frac{6}{8}.$$

Then, $E(X)=0=E(Y)$ and $E(X^2)=1=E(Y^2)$ but $E(X^4)=1\neq 4=E(Y^4)$. $\square$

Exercise 5.5.14. Prove the converse of Lemma 5.2.1. That is, prove that if $\{X_n\}$ converges to X almost surely, then for each $\epsilon>0$ we have $P(|X_n-X|\geq\epsilon\ i.o.)=0$.

Solution. From

$$1=P(\lim_n X_n=X)=P(\bigcap_{\epsilon>0}(\liminf_n |X_n-X|<\epsilon))=1-P(\bigcup_{\epsilon>0}(\limsup_n |X_n-X|\geq\epsilon)),$$

it follows:

$$P(\bigcup_{\epsilon>0}(\limsup_n |X_n-X|\geq\epsilon))=0.$$

On the other hand:

$$\forall\epsilon>0:(\limsup_n |X_n-X|\geq\epsilon)\subseteq(\bigcup_{\epsilon>0}(\limsup_n |X_n-X|\geq\epsilon)),$$

hence:

$$\forall\epsilon>0:P(\limsup_n |X_n-X|\geq\epsilon)=0.$$

$\square$

Chapter 6

Distributions of random variables

Exercise 6.3.2. Suppose $P(Z = 0) = P(Z = 1) = \frac{1}{2}$, that $Y \sim N(0, 1)$, and that Y and Z are independent. Set $X = YZ$. What is the law of X?

Solution. Using the definition of conditional probability given in Page 84, for any Borel Set $B \subseteq \mathbb{R}$ we have that:

$$\begin{aligned}
\mathcal{L}(X)(B) &= P(X \in B) \\
&= P(X \in B|Z = 0)P(Z = 0) + P(X \in B|Z = 1)P(Z = 1) \\
&= P(0 \in B)\frac{1}{2} + P(Y \in B)\frac{1}{2} \\
&= \frac{(\delta_0 + \mu_N)}{2}(B).
\end{aligned}$$

Therefore, $\mathcal{L}(X) = \frac{(\delta_0+\mu_N)}{2}$.□

Exercise 6.3.4. Compute $E(X), E(X^2)$, and $Var(X)$, where the law of X is given by

(a) $\mathcal{L}(X) = \frac{1}{2}\delta_1 + \frac{1}{2}\lambda$, where λ is Lebesgue measure on [0,1].
(b) $\mathcal{L}(X) = \frac{1}{3}\delta_2 + \frac{2}{3}\mu_N$, where μ_N is the standard normal distribution $N(0, 1)$.

Solution. Let $\mathcal{L}(X) = \sum_{i=1}^{n} \beta_i \mathcal{L}(X_i)$ where $\sum_{i=1}^{n} \beta_i = 1, 0 \leq \beta_i \leq 1$ for all $1 \leq i \leq n$. Then, for any Borel measurable function $f : \mathbb{R} \to \mathbb{R}$, combining Theorems 6.1.1, and 6.2.1 yields:

$$E_p(f(X)) = \sum_{i=1}^{n} \beta_i \int_{-\infty}^{\infty} f(t)\mathcal{L}(X_i)(dt).$$

Using the above result and considering $I(t) = t$, it follows:

(a)

$E_P(X) = \frac{1}{2}\int_{-\infty}^{\infty} I(t)\delta_1(dt) + \frac{1}{2}\int_{-\infty}^{\infty} I(t)\lambda(dt) = \frac{1}{2}(1) + \frac{1}{2}(\frac{1}{2}) = \frac{3}{4}$.
$E_P(X^2) = \frac{1}{2}\int_{-\infty}^{\infty} I^2(t)\delta_1(dt) + \frac{1}{2}\int_{-\infty}^{\infty} I^2(t)\lambda(dt) = \frac{1}{2}(1) + \frac{1}{2}(\frac{1}{3}) = \frac{2}{3}$.

$Var(X) = E_p(X^2) - E_p^2(X) = \frac{5}{48}$.

(b)

$E_P(X) = \frac{1}{3}\int_{-\infty}^{\infty} I(t)\delta_2(dt) + \frac{2}{3}\int_{-\infty}^{\infty} I(t)\mu_N(dt) = \frac{1}{3}(2) + \frac{2}{3}(0) = \frac{2}{3}$.
$E_P(X^2) = \frac{1}{3}\int_{-\infty}^{\infty} I^2(t)\delta_2(dt) + \frac{2}{3}\int_{-\infty}^{\infty} I^2(t)\mu_N(dt) = \frac{1}{3}(4) + \frac{2}{3}(1) = 2$.
$Var(X) = E_p(X^2) - E_p^2(X) = \frac{14}{9}$.□

Exercise 6.3.6. Let X and Y be random variables on some probability triple $(\Omega, \mathcal{F}, P)$. Suppose $E(X^4) < \infty$, and that $P(m \leq X \leq z) = P(m \leq Y \leq z)$ for all integers m and all $z \in \mathbb{R}$. Prove or disprove that we necessarily have $E(X^4) = E(Y^4)$.

Solution. Yes. First from $0 \leq P(X < m) \leq F_X(m), (m \in \mathbb{Z})$ and $\lim_{m\to-\infty} F_X(m) = 0$ it follows that:

$$\lim_{m\to-\infty} P(X < m) = 0.$$

Next, using recent result we have:

$$\begin{aligned} F_X(z) &= F_X(z) - \lim_{m\to-\infty} P(X < m) \\ &= \lim_{m\to-\infty} (P(X \leq z) - P(X < m)) \\ &= \lim_{m\to-\infty} P(m \leq X \leq z) \\ &= \lim_{m\to-\infty} P(m \leq Y \leq z) \\ &= F_Y(z). \end{aligned}$$

for all $z \in \mathbb{R}$.Therefore, by Proposition 6.0.2, $\mathcal{L}(X) = \mathcal{L}(Y)$ and by Corollary 6.1.3, the desired result follows.□

Exercise 6.3.8. Consider the statement : $f(x) = (f(x))^2$ for all $x \in \mathbb{R}$.
(a) Prove that the statement is true for all indicator functions $f = 1_B$.
(b) Prove that the statement is not true for the identity function $f(x) = x$.
(c) Why does this fact not contradict the method of proof of Theorem 6.1.1?

Solution. (a) $1_B^2 = 1_{B\cap B} = 1_B$, for all Borel measurable sets $B \subseteq \mathbb{R}$.

(b) $f(4) = 4 \neq 16 = (f(4))^2$.

(c) The main reason is the fact that the functional equation $f(x) = (f(x))^2$ is not stable when the satisfying function f is replaced by a linear combination such as $\sum_{i=1}^{n} a_i f_i$. Thus, in contrary to the method of Proof of Theorem 6.1.1, we cannot pass the stability of the given functional equation from indicator function to the simple function.□

Chapter 7

Stochastic processes and gambling games

Exercise 7.4.2. For the stochastic process $\{X_n\}$ given by (7.0.2), compute (for $n, k > 0$)
(a) $P(X_n = k)$.
(b) $P(X_n > 0)$.

Solution. (a) $P(X_n = k) = P(\sum_{i=1}^n r_i = \frac{n+k}{2}) = \binom{n}{\frac{n+k}{2}} 2^{-n}$ if $n + k = 2, 4, ..., 2n$, 0 etc.

(b) $P(X_n > 0) = \sum_{k=1}^n P(X_n = k) = \sum_{1 \le \frac{n+k}{2} \le n} \binom{n}{\frac{n+k}{2}} 2^{-n} = \sum_{u=\lfloor \frac{n}{2} \rfloor + 1}^n \binom{n}{u} 2^{-n}$. Using the binomial expansion $(1+1)^n = \sum_{k=0}^n \binom{n}{k}$ and the fact $\binom{n}{k} = \binom{n}{n-k}$ for all $0 \le k \le n$, we can further simplify the above result

$$
\begin{aligned}
P(X_n > 0) &= 2^{-n} \left(\sum_{u=0}^{n} \binom{n}{u} - \sum_{u=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{u} \right) \\
&= \begin{cases} 2^{-n} \left(\sum_{u=0}^n \binom{n}{u} - \frac{1}{2} \sum_{u=0}^n \binom{n}{u} \right) & \text{if } n \text{ is odd} \\ 2^{-n} \left(\sum_{u=0}^n \binom{n}{u} - \frac{\sum_{u=0}^n \binom{n}{u} - \binom{n}{\frac{n}{2}}}{2} \right) & \text{if } n \text{ is even} \end{cases} \\
&= \begin{cases} 2^{-n} \left(2^n - 2^{n-1} \right) & \text{if } n \text{ is odd} \\ 2^{-n} \left(2^n - 2^{n-1} + \frac{1}{2} \binom{n}{\frac{n}{2}} \right) & \text{if } n \text{ is even} \end{cases} \\
&= \begin{cases} \frac{1}{2} & \text{if } n \text{ is odd} \\ \frac{1}{2} + 2^{-n-1} \binom{n}{\frac{n}{2}} & \text{if } n \text{ is even} \end{cases} .
\end{aligned}
$$

Alternatively, since $P(X_n < 0) = P(X_n > 0)$, one can use instead

$$
P(X_n > 0) = \frac{1}{2}(1 - P(X_n = 0)),
$$

which gives:

- For n odd, $P(X_n = 0) = 0$, so $P(X_n > 0) = \frac{1}{2}$;
- For n even, $P(X_n = 0) = \binom{n}{\frac{n}{2}} 2^{-n}$, so $P(X_n > 0) = \frac{1}{2}(1 - \binom{n}{\frac{n}{2}} 2^{-n}) = \frac{1}{2} + 2^{-n-1} \binom{n}{\frac{n}{2}}$. $\square$

Exercise 7.4.4. For the gambler's ruin model of Subsection 7.2, with $c = 10,000$ and $p = 0.49$, find the smallest integer a such that $s_{c,p}(a) \geq \frac{1}{2}$. Interpret your result in plain English.

Solution. Substituting $p = 0.49$, $q = 0.51$, and $c = 10,000$ in the equation (7.2.2), and considering $s_{c,p}(a) \geq \frac{1}{2}$ it follows:

$$\frac{1 - (\frac{0.51}{0.49})^a}{1 - (\frac{0.51}{0.49})^{10,000}} \geq \frac{1}{2}$$

or

$$a \geq \frac{\ln(\frac{1}{2}((\frac{0.51}{0.49})^{10,000} - 1) + 1)}{\ln(\frac{51}{49})} \approx 9982.67$$

and the smallest positive integer value for a is 9983.
This result means that if we start with \$ 9983 (i.e. $a = 9983$) and our aim is to win \$ 10,000 before going broke (i.e. $c = 10,000$), then with the winning probability of %49 in each game (i.e.$p = 0.49$) our success probability, that we achieve our goal, is at least %50. $\square$

Exercise 7.4.6. Let W_n be i.i.d. with $P(W_n = 1) = P(W = 0) = \frac{1}{4}$ and $P(W_n = -1) = \frac{1}{2}$, and let a be a positive integer. Let $X_n = a + W_1 + W_2 + ... + W_n$, and let $\tau_0 = \inf\{n \geq 0; X_n = 0\}$. Compute $P(\tau_0 < \infty)$.

Solution. Let $0 \leq a \leq c$ and $\tau_c = \inf\{n \geq 0 : X_n = c\}$. Consider $s_c(a) = P(\tau_c < \tau_0)$,

$$\begin{aligned} s_c(a) &= P(\tau_c < \tau_0 | W_1 = 0)P(W_1 = 0) + P(\tau_c < \tau_0 | W_1 = 1)P(W_1 = 1) + P(\tau_c < \tau_0 | W_1 = -1)P(W_1 = -1) \\ &= \frac{1}{4}s_c(a) + \frac{1}{4}s_c(a+1) + \frac{1}{2}s_c(a-1), \end{aligned}$$

where $1 \leq a \leq c-1$, $s_c(0) = 0$, and $s_c(c) = 1$. Hence, $s_c(a+1) - s_c(a) = 2(s_c(a) - s_c(a-1))$ for all $1 \leq a \leq c-1$. Now, Solving this equation yields:

$$s_c(a) = \frac{2^a - 1}{2^c - 1} \qquad 0 \leq a \leq c.$$

On the other hand, $\{\tau_0 < \tau_c\} \nearrow \{\tau_0 < \infty\}$ if and only if $\{\tau_c \leq \tau_0\} \searrow \{\tau_0 = \infty\}$. Therefore:

$$\begin{aligned} P(\tau_0 < \infty) &= 1 - P(\tau_0 = \infty) \\ &= 1 - \lim_{c \to \infty} P(\tau_c \leq \tau_0) \\ &= 1 - \lim_{c \to \infty} S_c(a) \\ &= 1. \end{aligned}$$

$\square$

Exercise 7.4.8. In gambler's ruin, recall that $\{\tau_c < \tau_0\}$ is the event that the player eventually wins, and $\{\tau_0 < \tau_c\}$ is the event that the player eventually losses.
(a) Give a similar plain -English description of the complement of the union of these two events, i.e. $(\{\tau_c < \tau_0\} \cup \{\tau_0 < \tau_c\})^c$.
(b) Give three different proofs that the event described in part (a) has probability 0: one using Exercise 7.4.7; a second using Exercise 7.4.5; and a third recalling how the probabilities $s_{c,p}(a)$ were computed in the text, and seeing to what extent the computation would have differed if we had instead replaced $s_{c,p}(a)$ by $S_{c,p}(a) = P(\{\tau_c \leq \tau_0\})$.
(c) Prove that, if $c \geq 4$, then the event described in part (a) contains uncountably many outcomes(i.e. that uncountably many different sequences $Z_1, Z_2, ...$ correspond to this event, even though it has probability zero).

Solution. (a) The event $(\{\tau_c < \tau_0\} \cup \{\tau_0 < \tau_c\})^c = \{\tau_0 = \tau_c\}$ is the event that the player is both winner (winning $a - c$ dollar) and loser (losing a dollar) at the same time.

(b) Method (1):
$P(\{\tau_0 = \tau_c\}) = 1 - P(\{\tau_c < \tau_0\} \cup \{\tau_0 < \tau_c\}) = 1 - (r_{c,p}(a) + s_{c,p}(a)) = 0.$
Method (2):.
Method (3):
Let $S_{c,p}(a) = P(\{\tau_c \leq \tau_0\})$. Then, for $1 \leq a \leq c-1$ we have that:

$$S_{c,p}(a) = P(\{\tau_c \leq \tau_0\}|Z_1 = -1)P(Z_1 = -1) + P(\{\tau_c \leq \tau_0\}|Z_1 = 1)P(Z_1 = 1) = qS_{c,p}(a-1) + pS_{c,p}(a+1)$$

where $S_{c,p}(0) = 0$ and $S_{c,p}(c) = 1$. Solving the above equation, it follows $S_{c,p}(a) = s_{c,p}(a)$ for all $0 \leq a \leq c$. Thus,

$$P(\{\tau_c = \tau_0\}) = S_{c,p}(a) - s_{c,p}(a) = 0$$

for all $0 \leq a \leq c$.

(c) We prove the assertion for the case $a = 1$ and $c \geq 4$ (the case $a \geq 2$ and $c \geq 4$ is an straightforward generalization). Consider all sequences $\{Z_n\}_{n=1}^{\infty}$ of the form :

$$Z_1 = 1, Z_2 = \pm 1, Z_3 = -Z_2, Z_4 = \pm 1, Z_5 = -Z_4, ..., Z_{2n} = \pm 1, Z_{2n+1} = -Z_{2n}, ...$$

Since each Z_{2n} , $n = 1, 2, ...$ is selected in 2 ways, there are uncountably many sequences of this type (In fact there is an onto function f from the set of all of these sequences to the closed unite interval defined by

$$f(\{Z_n\}_{n=1}^{\infty}) = \sum_{n=1}^{\infty}(\frac{\text{sgn}(Z_{2n}) + 1}{2})2^{-n}).$$

In addition, a simple calculation shows that :

$$X_1 = 2, X_2 = 1 \text{ or } 3, X_3 = 2, X_4 = 1 \text{ or } 3, X_5 = 2, ..., X_{2n} = 1 \text{ or } 3, X_{2n+1} = 2, ...$$

□

Exercise 7.4.10. Consider the gambling policies model, with $p = \frac{1}{3}, a = 6$, and $c = 8$.
(a) Compute the probability $s_{c,p}(a)$ that the player will win (i.e. hit c before hitting 0) if they bet \$1 each time(i.e. if $W_n \equiv 1$).
(b) Compute the probability that the player will win if they bet \$ 2 each time (i.e. if $W_n \equiv 2$).
(c) Compute the probability that the player will win if they employ the strategy of Bold play(i.e., if $W_n = \min(X_{n-1}, c - X_{n-1})$).

Solution.(a) For $W_n = 1, p = \frac{1}{3}, a = 6, c = 8, q = \frac{2}{3}$ and $\frac{q}{p} = 2$, it follows:

$$s_{c,p}(a) = \frac{2^a - 1}{2^c - 1} \approx 0.247058823.$$

(b) For $W_n = 2, p = \frac{1}{3}, a = 3, c = 4, q = \frac{2}{3}$ and $\frac{q}{p} = 2$, it follows:

$$s_{c,p}(a) = \frac{2^a - 1}{2^c - 1} \approx 0.46666666.$$

(c) For $X_n = 6 + W_1Z_1 + W_2Z_2 + ... + W_nZ_n$, $W_n = \min(X_{n-1}, c - X_{n-1})$, $Z_n = \pm 1$ and $c = 8$ it follows:

$$W_1 = \min(X_0, 8 - X_0) = \min(6, 8 - 6) = 2,$$

$$W_2 = \min(X_1, 8 - X_1) = \min(6 + 2Z_1, 2 - 2Z_1) = 0 \text{ if } Z_1 = 1, 4 \text{ if } Z_1 = -1,$$

$$\begin{aligned} W_3 &= \min(X_2, 8 - X_2) \\ &= \min(6 + 2Z_1 + W_2Z_2, 2 - 2Z_1 - W_2Z_2) \\ &= 0\text{if}(Z_1 = 1, W_2 = 0), 0 \text{ if } (Z_1 = -1, W_2 = 4, Z_2 = 1), 8 \text{ if } (Z_1 = -1, W_2 = 4, Z_2 = -1), \end{aligned}$$

$$\begin{aligned} W_4 &= \min(X_3, 8 - X_3) \\ &= \min(6 + 2Z_1 + W_2Z_2 + W_3Z_3, 2 - 2Z_1 - W_2Z_2 - W_3Z_3) \\ &= 0 \ \text{ if } (Z_1 = 1 or Z_2 = 1), 0 \text{ if } (Z_1 = Z_2 = -1, Z_3 = 1), -8 \text{ if } (Z_1 = Z_2 = Z_3 = -1), \end{aligned}$$

When $Z_1 = Z_2 = -1$, the event τ_0 occurs. Hence:

$$\begin{aligned} P(\tau_c < \tau_0) &= P(Z_1 = 1) + P(Z_1 = -1, Z_2 = 1) \\ &= P(Z_1 = 1) + P(Z_1 = -1)P(Z_2 = 1) \\ &= \frac{1}{3} + \frac{2}{3}\frac{1}{3} \\ &= 0.5555555. \end{aligned}$$

□

Chapter 8

Discrete Markov chains

Exercise 8.5.2. For any $\epsilon > 0$, give an example of an irreducible Markov chain on a countably infinite state space , such that $|p_{ij} - p_{ik}| \leq \epsilon$ for all states i, j, and k.

Solution. Given $\epsilon > 0$. If $\epsilon \geq 1$, then put $S = \mathbb{N}, v_i = 2^{-i}(i \in \mathbb{N})$, and $p_{ij} = 2^{-j}(i, j \in \mathbb{N})$, giving:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2^2} & \frac{1}{2^3} & \cdot & \cdot & \cdot \\ \frac{1}{2} & \frac{1}{2^2} & \frac{1}{2^3} & \cdot & \cdot & \cdot \\ \frac{1}{2} & \frac{1}{2^2} & \frac{1}{2^3} & \cdot & \cdot & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

If $0 < \epsilon < 1$, then put $S = \mathbb{N}, v_i = 2^{-i}(i \in \mathbb{N})$. Define $n_0 = \max\{n \in \mathbb{N} : n\epsilon < 1\}$ and put $p_{ij} = \epsilon$ if $(i = 1, 2, ..., j = 1, 2, ..., n_0)$, $(1 - n_0\epsilon)2^{-(j-n_0)}$ if $(i = 1, 2, ..., j = n_0 + 1, ...)$, giving:

$$P = \begin{pmatrix} \epsilon & \cdots & \epsilon & \frac{1-n_0\epsilon}{2} & \frac{1-n_0\epsilon}{2^2} & \frac{1-n_0\epsilon}{2^3} & \cdots \\ \epsilon & \cdots & \epsilon & \frac{1-n_0\epsilon}{2} & \frac{1-n_0\epsilon}{2^2} & \frac{1-n_0\epsilon}{2^3} & \cdots \\ \epsilon & \cdots & \epsilon & \frac{1-n_0\epsilon}{2} & \frac{1-n_0\epsilon}{2^2} & \frac{1-n_0\epsilon}{2^3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

In both cases, $|p_{ij} - p_{ik}| \leq \epsilon$ for all states i, j, and k (Check!). In addition, since $P_{ij} > 0$ for all $i, j \in \mathbb{N}$, it follows that the corresponding Markov chain in the Theorem 8.1.1., is irreducible.

Note: A minor change in above solution shows that, in fact, there are uncountably many Markov chains having this property (Check!). □

Exercise 8.5.4. Given Markov chain transition probabilities $\{P_{ij}\}_{i,j \in S}$ on a state space S, call a subset $C \subseteq S$ closed if $\sum_{j \in C} p_{ij} = 1$ for each $i \in C$. Prove that a Markov chain is irreducible if and only if it has no closed subsets (aside from the empty set and S itself).

Solution. First, let the given Markov chain has a proper closed subset C. Since $\sum_{j \in S} p_{ij} = 1$ for each $i \in S$ and $\sum_{j \in C} p_{ij} = 1$ for each $i \in C$, we conclude that :

$$\sum_{j \in S-C} p_{ij} = \sum_{j \in S} p_{ij} - \sum_{j \in C} p_{ij} = 0, (i \in C).$$

Consequently, $p_{ij} = 0$ for all $i \in C$ and $j \in S - C$. Next, consider the Chapman-Kolmogorov equation:

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}, (1 \leq k \leq n, n \in \mathbb{N}).$$

Specially, for $k = n - 1, j \in S - C$, and $i \in C$, applying the above result gives:

$$p_{ij}^{(n)} = \sum_{r \in S - C} p_{ir}^{(n-1)} p_{rj}, (n \in \mathbb{N}).$$

Thus, the recent equation, inductively, yields:

$$p_{ij}^{(n)} = 0 (i \in C, j \in S - C, n \in \mathbb{N}),$$

where $C \neq \emptyset$ and $S - C \neq \emptyset$. Therefore, the given Markov chain is reducible.

Second, assume the given Markov chain is reducible. Hence, there are $i_0, j_0 \in S$ such that $p_{i_0 j_0}^{(n)} = 0$for all $n \in \mathbb{N}$. On the other hand, $C \subseteq S$ is closed if and only if for all $i \in C$ and $j \in S$ if $p_{ij}^{(n_0)} = 0$ for some $n_0 \in \mathbb{N}$, then $j \in C$

(Proof. Let $C \subseteq S$ be closed. If $i \in C$ and $j \in S - C$ with $p_{ij}^{(n_0)} > 0$ for some $n_0 \in \mathbb{N}$, then $\sum_{k \in C} p_{ik}^{(n_0)} < 1$, a contradiction. Conversely, if the condition is satisfied and C is not closed, then $\sum_{j \in C} p_{ij} < 1$, for some $i \in C$; hence $p_{ij} > 0$ for some $i \in C$ and $j \in S - C$, a contradiction.).

Now, it is sufficient to take $C = S - \{j_0\}$. $\square$

Exercise 8.5.6. Consider the Markov chain with state space $S = \{1, 2, 3\}$ and transition probabilities $p_{12} = p_{23} = p_{31} = 1$. Let $\pi_1 = \pi_2 = \pi_3 = \frac{1}{3}$.
(a) Determine whether or not the chain is irreducible.
(b) Determine whether or not the chain is aperiodic.
(c) Determine whether or not the chain is reversible with respect to $\{\pi_i\}$.
(d) Determine whether or not $\{\pi_i\}$ is a stationary distribution.
(e) Determine whether or not $\lim_{n \to \infty} p_{11}^{(n)} = \pi_1$.

Solution. (a) Yes. A simple calculation shows that for any nonempty subset $C \subsetneq S$, there is $i \in C$ such that $\sum_{j \in C} p_{ij} < 1$, (Check!) . Thus, using Exercise 8.5.4., the given Markov chain is irreducible.

(b) No. Since the given Markov chain is irreducible, by Corollary 8.3.7., all of its states have the same period. Hence, $period(1) = period(2) = period(3)$. Let $i = 1$ and consider the Chapman-Kolmogorov equation in the solution of Exercise 8.5.4. Then:

$$\begin{aligned} p_{11}^{(n)} &= \sum_{r=1}^{3} P_{1r}^{(n-1)} p_{r1} = p_{13}^{(n-1)}, \\ p_{13}^{(n)} &= \sum_{r=1}^{3} P_{1r}^{(n-2)} p_{r3} = p_{12}^{(n-2)}, \\ p_{12}^{(n)} &= \sum_{r=1}^{3} P_{1r}^{(n-3)} p_{r2} = p_{11}^{(n-3)}, \end{aligned}$$

besides:

$$\begin{aligned} p_{11}^{(1)} &= 0, \\ p_{11}^{(2)} &= \sum_{i_{k+1}=1}^{3} P_{1i_{k+1}} p_{i_{k+1}1} = 0, \\ p_{11}^{(3)} &= \sum_{i_{k+1}=1}^{3} \sum_{i_{k+2}=1}^{3} P_{1i_{k+1}} P_{i_{k+1}i_{k+2}} p_{i_{k+2}1} = 1, \end{aligned}$$

implying:

$$p_{11}^{(n)} = 1 \quad \text{if} \quad 3|n, 0 \quad \text{etc.}$$

Consequently, $period(1) = 3 \neq 1$.

(c) No. Let for all $i, j \in S$, $\pi_i p_{ij} = \pi_j p_{ji}$. Since, $\pi_i = \pi_j = \frac{1}{3}$, it follows that for all $i, j \in S$, $p_{ij} = p_{ji}$. On the other hand, $p_{12} = 1 \neq 0 = p_{21}$, showing that this chain is not reversible with respect to $\{\pi_i\}_{i=1}^{3}$.

(d) Yes. Since $\sum_{i\in S} p_{ij} = 1$ for any $i, j \in S$ and $\pi_i = \pi_j = \frac{1}{3}$, it follows that for any $j \in S$, $\sum_{i\in S} \pi_i p_{ij} = \frac{1}{3}\sum_{i\in S} p_{ij} = \frac{1}{3} = \pi_j$.

(e) No. Since $p_{11}^{(n)} = 1 \quad \text{if} \quad 3|n, 0 \quad$ etc., the given limit does not exist.□

Exercise 8.5.8. Prove the identity $f_{ij} = p_{ij} + \sum_{k\neq j} p_{ik} f_{kj}$.

Solution.

$$\begin{aligned} f_{ij} &= f_{ij}^{(1)} + \sum_{n=2}^{\infty} f_{ij}^{(n)} \\ &= p_{ij} + \sum_{n=2}^{\infty} \sum_{k\neq j} P_i(X_1 = k, X_2 \neq j, ..., X_{n-1} \neq j, X_n = j) \\ &= p_{ij} + \sum_{k\neq j} \sum_{n=2}^{\infty} P_i(X_1 = k, X_2 \neq j, ..., X_{n-1} \neq j, X_n = j) \\ &= p_{ij} + \sum_{k\neq j} \sum_{n=2}^{\infty} P_i(X_1 = k) P_i(X_2 \neq j, ..., X_{n-1} \neq j, X_n = j) \\ &= p_{ij} + \sum_{k\neq j} p_{ik} \sum_{n=2}^{\infty} P_i(X_2 \neq j, ..., X_{n-1} \neq j, X_n = j) \\ &= p_{ij} + \sum_{k\neq j} p_{ik} f_{kj}. \end{aligned}$$

□ **Exercise 8.5.10.** Consider a Markov chain (not necessarily irreducible) on a finite state space.
(a) Prove that at least one state must be recurrent.
(b) Give an example where exactly one state is recurrent (and all the rest are transient).
(c) Show by example that if the state space is countably infinite then part (a) is no longer true.

Solution. (a) To apply the Kolmogorov 0-1 law, one can consider the stopping time $\tau_n = \{n > \tau_{n-1} : X_n = i_0, \text{ but } X_m \neq i_0 \text{ for } \tau_{n-1} < m < \tau_m\}$. Then, $\{\tau_{n+1} - \tau_n\}$ are independent by the strong Markov property. Moreover, $\{X_n = i_0 \text{ i.o}\}$ lies in the tail σ-algebra of $\{\tau_{n+1} - \tau_n\}$. It follows from the Kolmogorov 0-1 law that $\mathbf{P}_{i_0}(X_n = i_0 \text{ i.o.}) = 1$. However, the strong Markov property is not included in the textbook
Notation: we say state j is accessible from i if there exists $n > 0$ such that $\mathbf{P}_i(X_n = j) > 0$, denoted by $i \to j$. On the other hand, j is not accessible from i if $\mathbf{P}_i(X_n = j) = 0$ for all $n > 0$, denoted by $i \not\to j$.
Assume all the states are transient. Consider an arbitrary state $i_1 \in \mathcal{S}$. Since i_1 is transient, $\mathbf{P}_{i_1}(X_n = i_1 \text{ i.o.}) = 0$ by Theorem 8.2.1. It follows that there exists $i_2 \in \mathcal{S}$ and $i_2 \neq i_1$ such that $i_1 \to i_2$ and $i_2 \not\to i_1$.
Similarly, since i_2 is transient, there exists $i_3 \in \mathcal{S}$ and $i_3 \neq i_2$, $i_3, \neq i_1$ such that $i_2 \to i_3$ and $i_3 \not\to i_2$. Since $i_1 \to i_2$ and $i_2 \to i_3$, we have $i_1 \to i_3$. Moreover, $i_3 \not\to i_1$. Indeed, assume $i_3 \to i_1$, since $i_1 \to i_2$, we can get $i_3 \to i_2$, contradiction.
Continue in this way, the process will stop at i_n, where $n = |\mathcal{S}| < \infty$, since we only have n distinct states. We then have $i_1 \to i_k$ and $i_k \not\to i_1$. By the construction, $i_k \not\to i_j$ for $j < k$, i.e. $\mathbf{P}_{i_k}(X_n \in \mathcal{S}\backslash\{i_k\}) = 0$. It follows that $\mathbf{P}_{i_k}(X_n = i_k) = 1$, which implies i_k is recurrent. Contradiction.
(b) Let $S = \{1, 2\}$ and $(p_{ij}) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. Using Chapman-Kolmogorov equation:

$$p_{i2}^{(n)} = \sum_{r=1}^{2} p_{ir}^{(n-1)} p_{r2} = 0(i = 1, 2, n = 1, 2, 3, ...).$$

Specially, $p_{22}^{(n)} = 0 (n \in \mathbb{N})$, and, hence, $\sum_{n=1}^{\infty} p_{22}^{(n)} = 0 < \infty$. Thus, by Theorem 8.2.1. , it follows that the state 2 is transient . Eventually, by part (a), the only remaining state, which is 1, is recurrent.

(c) Any state $i \in \mathbb{Z}$ in the simple asymmetric random walk is not recurrent (see page 87). □

Exercise 8.5.12. Let $P = (p_{ij})$ be the matrix of transition probabilities for a Markov chain on a finite state space.
(a) Prove that P always has 1 as an eigenvalue.
(b) Suppose that v is a row eigenvector for P corresponding to the eigenvalue 1, so that $vP = v$. Does v necessarily correspond to a stationary distribution? Why or why not?

Solution. (a) Let $|S| = n$ and $P = (p_{ij})$. Since $|S| < \infty$, without loss of generality we can assume $S = \{1, 2, ..., n\}$. Consider

$$[\mu^{(0)}]^t = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Then:

$$P[\mu^{(0)}]^t = \begin{pmatrix} \sum_{j=1}^{n} p_{1j}.1 \\ \sum_{j=1}^{n} p_{1j}.1 \\ \vdots \\ \sum_{j=1}^{n} p_{1j}.1 \end{pmatrix} = [\mu^{(0)}]^t.$$

So, $\lambda = 1$ is an eigenvalue for P.

(b) Generally No. As a counterexample, we can consider $S = \{1, 2, ..., n\}, P = I_{n\times n}$ and $v = (-\frac{1}{n})_{i=1}^{n}$ which trivially does not correspond to any stationary distribution.□

Exercise 8.5.14. Give an example of a Markov chain on a finite state space, such that three of the states each have a different period.

Solution. Consider the corresponding Markov chain of Theorem 8.1.1. to the state space $S = \{1, 2, ..., 6\}$ and the transition matrix:

$$(p_{ij}) = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \end{pmatrix}.$$

Then, since $p_{11} > 0, p_{23}p_{32} > 0, \text{and}, p_{45}p_{56}p_{64} > 0$, it follows $period(1) = 1, Period(2) = Period(3) = 2$, and $Period(4) = Period(5) = Period(6) = 3$, respectively .$\square$

Exercise 8.5.16. Consider the Markov chain with state space $S = \{1, 2, 3\}$ and transition probabilities given by :

$$(p_{ij}) = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/4 & 0 & 3/4 \\ 4/5 & 1/5 & 0 \end{pmatrix}.$$

(a) Find an explicit formula for $P_1(\tau_1 = n)$ for each $n \in \mathbb{N}$, where $\tau_1 = \inf\{n \geq 1 : X_n = 1\}$.
(b) Compute the mean return time $m_1 = E_1(\tau_1)$.
(c) Prove that this Markov chain has a unique stationary distribution, to be called $\{\pi_i\}$.
(d) Compute the stationary probability π_1.

Solution. (a) Let $a_n = P_1(\tau_1 = n)$. Computing a_n for $n = 1, 2, ..., 7$ yields:

$$\begin{aligned} a_1 &= 0 \\ a_2 &= p_{12}p_{21} + p_{13}p_{31} \\ a_3 &= p_{12}p_{23}p_{31} + p_{13}p_{32}p_{21} \\ a_4 &= p_{12}p_{23}p_{32}p_{21} + p_{13}p_{32}p_{23}p_{31} \\ a_5 &= p_{13}p_{32}p_{23}p_{32}p_{21} + p_{12}p_{23}p_{32}p_{23}p_{31} \\ a_6 &= p_{13}p_{32}p_{23}p_{32}p_{23}p_{31} + p_{12}p_{23}p_{32}p_{23}p_{32}p_{21} \\ a_7 &= p_{13}p_{32}p_{23}p_{32}p_{23}p_{32}p_{21} + p_{12}p_{23}p_{32}p_{23}p_{32}p_{23}p_{31}. \end{aligned}$$

In general, it follows by induction (Check!) that:

$$\begin{aligned} a_{2m} &= (p_{23}p_{32})^{m-1}(p_{12}p_{21} + p_{13}p_{31}) \qquad (m \in \mathbb{N}) \\ a_{2m+1} &= (p_{23}p_{32})^{m-1}(p_{12}p_{23}p_{31} + p_{13}p_{32}p_{21}) \qquad (m \in \mathbb{N}). \end{aligned}$$

(b) Using part (a) we conclude:

$$\begin{aligned}
m_1 &= E(\tau_1)\\
&= \sum_{n=1}^{\infty} nP_1(\tau_1 = n)\\
&= \sum_{m=1}^{\infty} 2mP_1(\tau_1 = 2m) + \sum_{m=1}^{\infty}(2m+1)P_1(\tau_1 = 2m+1)\\
&= 2(p_{12}p_{21} + p_{13}p_{31})\sum_{m=1}^{\infty} m(p_{23}p_{32})^{m-1} + 2(p_{12}p_{23}p_{31} + p_{13}p_{32}p_{21})\sum_{m=1}^{\infty}(2m+1)(p_{23}p_{32})^{m-1}\\
&= \frac{2(p_{12}p_{21} + p_{13}p_{31})}{(1-p_{23}p_{32})^2} + (p_{12}p_{23}p_{31} + p_{13}p_{32}p_{21})(\frac{2}{(1-p_{23}p_{32})^2} + \frac{1}{(1-p_{23}p_{32})}).
\end{aligned}$$

(c) First, we show that the given Markov chain is irreducible. Using Chapman-Kolmogorov equation:

$$p_{ij}^{(n)} = \sum_{r=1}^{3} p_{ir}p_{rj}^{(n-1)} \qquad i,j = 1,2,3, n = 1,2,...,$$

it follows that for any $i, j = 1, 2, 3$, there is $n_0 \in \mathbb{N}$ such that $p_{ij}^{(n_0)} > 0$. In addition, a computation similar to part (a) shows that $m_i = E(\tau_i) < \infty$ for $i = 2, 3$. Hence, by Theorem 8.4.1., this Markov chain has a unique distribution $\{\pi_i\}_{i=1}^{3}$ given by $\pi_i = \frac{1}{m_i}$ for $i = 1, 2, 3$.

(d)

$$\pi_1 = \frac{1}{m_1} = \frac{(1-p_{23}p_{32})^2}{2(p_{12}p_{21} + p_{13}p_{31}) + (p_{12}p_{23}p_{31} + p_{13}p_{32}p_{21})(3 - p_{23}p_{32})}.$$

□

Exercise 8.5.18. Prove that if $f_{ij} > 0$ and $f_{ji} = 0$, then i is transient.

Solution. The comment by David said he could not quite understand the original proof (neither do I)
Since $f_{ij} > 0$, there exists $n \in \mathbb{N}$ such that $f_{ij}^{(n)} > 0$. Moreover, we have $f_{ji} = 0$ by assumption, which means the chain will not go back to i after reaching the state j. Hence,

$$\mathbf{P}_i(X_n = i \text{ for finitely many of } n) \geq f_{ij}^{(n)} > 0.$$

It follows that

$$\mathbf{P}_i(X_n = i \text{ i.o.}) = 1 - \mathbf{P}_i(X_n = i \text{ for finitely many of } n) \leq 1 - f_{ij}^{(n)}.$$

According to Theorem 8.2.1, i is transient.□

Exercise 8.5.20. (a) Give an example of a Markov chain on a finite state space which has multiple (i.e. two or more) stationary distributions.
(b) Give an example of a reducible Markov chain on a finite state space, which nevertheless has a unique stationary distribution.
(c) Suppose that a Markov chain on a finite state space is decomposable, meaning that the state space can be partitioned as $S = S_1 \cup S_2$, with S_i nonempty, such that $f_{ij} = f_{ji} = 0$ whenever $i \in S_1$ and $j \in S_2$. Prove that the chain has multiple stationary distribution.
(d) Prove that for a Markov chain as in part (b) , some states are transient.

Solution. (a) Consider a Markov chain with state space $S = \{i\}_{i=1}^{n}$ and transition matrix $(p_{ij}) = I_{n\times n}$ where $n \geq 3$. Then, any distribution $(\pi_i)_{i=1}^{n}$ with $\sum_{i=1}^{n} \pi_i = 1$ and $\pi_i \geq 0$ is its stationary distribution.

(b) Take $S = \{1, 2\}$ and:

$$(p_{ij}) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Applying Chapman-Kolmogorov equation:

$$p_{11}^{(n)} = \sum_{r=1}^{2} p_{1r}^{(n-1)} p_{r1} (n \in \mathbb{N}),$$

it follows that $p_{11}^{(n)} = 0 \qquad (n \in \mathbb{N})$ yielding that this chain is reducible . Let $(\pi_i)_{i=1}^{2}$ with $\pi_1 + \pi_2 = 1$ and $0 \leq \pi_1, \pi_2 \leq 1$ be a stationary distribution. Then, $\pi_1 = \pi_1 p_{11} + \pi_2 p_{21} = 0$ and $\pi_2 = 1 - \pi_1 = 1$. Thus, this chain has only one stationary distribution.

(c) Let $S = S_1 \cup S_2$ and $f_{ij} = 0 = f_{ji}$ for any $i \in S_1, j \in S_2$. Hence, $f_{ij} > 0$ if either $i, j \in S_1$ or $i, j \in S_2$. Therefore, the restriction of the given Markov chain on the state spaces $S_r (r = 1, 2)$ is irreducible. Hence, by Proposition 8.4.10, all the states of the state spaces $S_r (r = 1, 2)$ are positive recurrent. Now, by Theorem 8.4.9, there are unique stationary distributions for $S_r (r = 1, 2)$, say $(\pi_i)_{i=1}^{|S_1|}$ and $(\pi_i)_{i=|S_1|+1}^{|S_1|+|S_2|}$, respectively. Eventually, pick any $0 < \alpha < 1$ and consider the stationary distribution $(\pi_i)_{i=1}^{|S|}$ defined by:

$$\pi_i = \alpha \pi_i 1_{\{1,\ldots,|S_1|\}}(i) + (1 - \alpha) \pi_i 1_{\{|S_1|+1,\ldots,|S|\}}(i).$$

Note: This is not quite correct as stated, since the chain might not be irreducible even on the state soace S_r $(r = 1, 2)$. But it is still true that there is at least one stationary distribution on each S_r, hence at least two stationary distributions in total.

(d) Consider the solution of part (b). In that example, since $p_{21} = 0$ and $p_{21}^{(n)} = \sum_{r=1}^{2} p_{2r}^{(n-1)} p_{r1} = 0 \qquad (n \geq 2)$, it follows that $f_{21} = 0$. On the other hand, $p_{12} > 0$ and hence, $f_{12} > 0$. Now, by Exercise 8.5.18. the state $i = 1$ is transient. □

Chapter 9

More probability theorems

Exercise 9.5.2. Give an example of a sequence of Random variables which is unbounded but still uniformly integrable. For bonus points, make the sequence also be undominated , i.e. violate the hypothesis of the Dominated Convergence Theorem.

Solution. Let $\Omega = \mathbb{N}$, and $P(\omega) = 2^{-\omega}(\omega \in \Omega)$. For $n \in \mathbb{N}$, define:

$$\begin{aligned} X_n &: \Omega \to \mathbb{R} \\ X_n(\omega) &= n\delta_{\omega n}. \end{aligned}$$

Since $\lim_{n\to\infty} X_n(n) = \infty$, there is no $K > 0$ such that $|X_n| < K$ for all $n \in \mathbb{N}$. On the other hand, for any α there exists some $N_\alpha = \lceil \alpha \rceil$ such that :

$$|X_n| 1_{|X_n| \geq \alpha}(\omega) = X_n(\omega) \qquad n = N_\alpha, N_\alpha + 1, \cdots.$$

Thus, we have that:

$$\sup_n E(|X_n| 1_{|X_n|\geq\alpha}) = \sup_{n\geq N_\alpha} E(X_n) = \sup_{n\geq N_\alpha} (n2^{-n}) = N_\alpha 2^{-N_\alpha},$$

implying:

$$\lim_\alpha \sup_n E(|X_n| 1_{|X_n|\geq\alpha}) = \lim_\alpha N_\alpha 2^{-N_\alpha} = 0.$$

To get the bonus point, consider $P(\omega) = \frac{6}{\pi^2}\frac{1}{\omega^2}$.
Claim. $X_n(\omega)$ is undominated.
Suppose there exists a random variable $Y(\omega)$ such that $X_n(\omega) \leq Y(\omega)$ for all n, ω and $\mathbf{E}(Y) < \infty$. Since $X_n(\omega) \leq Y(\omega)$ for all $n, \omega \in \mathbb{N}$, $Y \geq \omega$ for all $\omega \in \mathbb{N}$ by the definition of X_n. It follows that

$$\sum_{\omega\in\mathbb{N}} \omega P(\omega) \leq \sum_{\omega\in\mathbb{N}} Y(\omega)P(\omega) \Rightarrow \mathbf{E}(Y) \geq \infty.$$

Contradiction. Therefore, $X_n(\omega)$ is undominated. □
To prove uniform integrability, we follow the same steps as before: For any α, there exists $N_\alpha = \lceil \alpha \rceil$ such that

$$|X_n| \mathbf{1}_{|X_n|\geq\alpha}(\omega) = X_n(\omega), \quad \forall n = N_\alpha, N_\alpha + 1, \ldots.$$

Then, we have

$$\sup_n \mathbf{E}(|X_n| \mathbf{1}_{|X_n|\geq\alpha}) = \sup_{n\geq N_\alpha} \mathbf{E}(X_n) = \sup_{n\geq N_\alpha} \left(n\frac{6}{\pi^2}\frac{1}{n^2} \right) = \frac{6}{\pi^2}\frac{1}{N_\alpha}$$

implying

$$\lim_\alpha \sup_n \mathbf{E}(|X_n| \mathbf{1}_{|X_n|\geq\alpha}) = \lim_\alpha \frac{6}{\pi^2}\frac{1}{N_\alpha} = 0.$$

□

Exercise 9.5.4. Suppose that $\lim_{n\to\infty} X_n(\omega) = 0$ for all $\omega \in \Omega$, but $\lim_{n\to\infty} E(X_n) \neq 0$. Prove that $E(\sup_n |X_n|) = \infty$.

Solution. Assume $E(\sup_n |X_n|) < \infty$. Take $Y = \sup_n |X_n|$. Now, according to the Theorem 9.1.2,

$$\lim_{n\to\infty} E(X_n) = E(\lim_{n\to\infty} X_n) = 0,$$

a contradiction.□

Exercise 9.5.6. Prove that Theorem 9.1.6. implies Theorem 9.1.2.

Solution. Let the assumptions of the Theorem 9.1.2. hold. Since $|X_n| \leq Y$, it follows $|X_n| 1_{|X_n|\geq\alpha} \leq Y 1_{Y\geq\alpha}$, and consequently , by taking first expectation from both sides of the inequality and then taking supremum of the result inequality , we have:

$$\sup_n E(|X_n| 1_{|X_n|\geq\alpha}) \leq E(Y 1_{Y\geq\alpha}) \qquad (\alpha > 0)$$

On the other hand, $\lim_\alpha E(Y 1_{Y\geq\alpha}) = 0$. Thus, by above inequality it follows

$$\lim_\alpha \sup_n E(|X_n| 1_{|X_n|\geq\alpha}) = 0.$$

Eventually, by Theorem 9.1.6, $\lim_n E(X_n) = E(X)$. □

Exercise 9.5.8. Let $\Omega = \{1, 2\}$, with $P(\{1\}) = P(\{2\}) = \frac{1}{2}$, and let $F_t(\{1\}) = t^2$, and $F_t(\{2\}) = t^4$,for $0 < t < 1$.
(a) What does Proposition 9.2.1. conclude in this case?
(b) In light of the above, what rule from calculus is implied by Proposition 9.2.1?

Solution. (a) Since :

$$E(F_t) = F_t(\{1\})P(\{1\}) + F_t(\{2\})P(\{2\}) = \frac{t^2 + t^4}{2} \leq 1 < \infty,$$

for all$0 < t < 1$ and, $F_t^{'}(\{1\}) = 2t$ and $F_t^{'}(\{2\}) = 4t^3$ exist for all $0 < t < 1$, by Theorem 9.2.1. it follows that $F_t^{'}$ is random variable. Besides, since $|F_t^{'}| \leq 4$ for all $0 < t < 1$ and $E(4) = 4 < \infty$, according to Theorem 9.2.1. $\phi(t) = E(F_t)$ is differentiable with finite derivative $\phi^{'}(t) = E(F_t^{'})$ for all $0 < t < 1$. Thus, $E(F_t^{'}) = t + 2t^3$ for all $0 < t < 1$.

(b) $\frac{d}{dx} \int^x f(t)dt = \int^x \frac{d}{dt} f(t)dt$. □

Exercise 9.5.10. Let $X_1, X_2, \cdots$, be i.i.d., each having the standard normal distribution $N(0, 1)$. Use Theorem 9.3.4. to obtain an exponentially decreasing upper bound on $P(\frac{1}{n}(X_1 + \cdots + X_n) \geq 0.1)$.

Solution. In this case, for the assumptions of the Theorem 9.3.4., $M_{X_i}(s) = \exp(\frac{s^2}{2}) < \infty$ for all $-a < s < b$ and all $a, b, > 0$, $m = 0$, and $\epsilon = 0.1$. Thus:

$$P(\frac{1}{n}(X_1 + \cdots + X_n) \geq 0.1) \leq \rho^n$$

where $\rho = \inf_{0<s<b}(\exp(\frac{s^2}{2} - 0.1s))$ for all $n \in \mathbb{N}$. Put $g(s) = \exp(\frac{s^2}{2} - 0.1s)$ for $0 < s < \infty$. Then, a simple calculation shows that for $b = 0.1$ the function g attains its infimum. Accordingly, $\rho = g(b) = 0.995012449$. □

Exercise 9.5.12. Let $\alpha > 2$, and let $M(t) = \exp(-|t|^\alpha)$ for $t \in \mathbb{R}$. Prove that $M(t)$ is not a moment generating function of any probability distribution.

Solution. Calculating the first two derivatives, we have:

$$\begin{aligned} M^{'}(t) &= -\alpha t|t|^{\alpha-2}\exp(-|t|^\alpha) \qquad -\infty < t < \infty \\ M^{''}(t) &= -\alpha|t|^{\alpha-2}((\alpha-1)-\alpha|t|^\alpha)\exp(-|t|^\alpha) \quad -\infty < t < \infty, \end{aligned}$$

yielding, $M^{'}(0) = 0$ and $M^{''}(0) = 0$. Now, assume $M = M_X$ for some random variable X. Then, applying the Remark on page 108 (i.e. $M^r(0) = E(X^r)$: $r \in \mathbb{N},$) for the case $r = 1, 2$, it follows $E(X^2) = E(X) = 0$, and, therefore, $Var(X) = 0$. Next, by Proposition 5.1.2, $P(|X| \geq \alpha) \leq \frac{Var(X)}{\alpha^2} = 0$ for all $\alpha > 0$, implying $P(X \neq 0) = 0$ or equivalently, $X \sim \delta_0$. Accordingly, $M_X(t) = 1$ for all $-\infty < t < \infty$, a clear contradiction. □

Exercise 9.5.14. Let λ be Lebesgue measure on $[0, 1]$, and let $f(x, y) = 8xy(x^2-y^2)(x^2+y^2)^{-3}$ for $(x, y) \neq (0, 0)$, with $f(0, 0) = 0$.
(a) Compute $\int_0^1(\int_0^1 f(x, y)\lambda(dy))\lambda(dx)$.
(b) Compute $\int_0^1(\int_0^1 f(x, y)\lambda(dx))\lambda(dy)$.
(c) Why does the result not contradict Fubini's Theorem.

Solution. (a) Take $u = y^2 + x^2$, $v = x^2$, then $du = 2ydy$, $u - 2x^2 = y^2 - x^2$ and $dv = 2xdx$. Accordingly:

$$\begin{aligned} \int_0^1(\int_0^1 f(x,y)dy)dx &= \int_0^1(\int_0^1 \frac{8xy(x^2-y^2)}{(x^2+y^2)^3}dy)dx \\ &= \int_0^1(-8x\int_0^1 \frac{y(y^2-x^2)}{(x^2+y^2)^3}dy)dx \\ &= \int_0^1(-8x(\frac{-y^2}{2(y^2+x^2)^2})|_0^1)dx \\ &= \int_0^1 \frac{4x}{(1+x^2)^2}dx \\ &= \int_0^1 \frac{2dv}{(1+v^2)^2} \\ &= 1. \end{aligned}$$

(b) Using integration by parts for the inside integral, it follows:

$$\begin{aligned}
\int_0^1 (\int_0^1 f(x,y)dx)dy &= \int_0^1 (\int_0^1 8xy(x^2-y^2)(x^2+y^2)^{-3}dx)dy \\
&= \int_0^1 (\int_0^1 4x(x^2-y^2)(x^2+y^2)^{-3}dx)2ydy \\
&= \int_0^1 (\frac{-2}{(y^2+1)^2})2ydy \\
&= -1.
\end{aligned}$$

(c) In this case, the hypothesis of the Fubini's Theorem is violated. In fact, by considering suitable Riemann sums for double integrals, it follows that $\int\int f^+\lambda(dx \times dy) = \int\int f^-\lambda(dx \times dy) = \infty$. □

Exercise 9.5.16. Let $X \sim N(a, v)$ and $Y \sim N(b, w)$ be independent. Let $Z = X + Y$. Use the convolution formulae to prove that $Z \sim N(a + b, v + w)$.

Solution. Let $a = b = 0$. we claim that if $X \sim N(0, v)$ and $Y \sim N(0, w)$, then $Z \sim N(0, v + w)$. To prove it, using convolution formulae, it follows:

$$\begin{aligned}
h_Z(z) &= \int_{-\infty}^{\infty} f_X(z-y)g_Y(y)dy \\
&= \frac{1}{\sqrt{2\pi v}}\frac{1}{\sqrt{2\pi w}}\int_{-\infty}^{\infty} \exp(-\frac{1}{2}(\frac{(z-y)^2}{v}+\frac{y^2}{w}))dy \\
&= \frac{\exp(-\frac{z^2}{v})}{\sqrt{2\pi v}\sqrt{2\pi w}}\int_{-\infty}^{\infty} \exp(-\frac{1}{2}((\frac{1}{w}+\frac{1}{v})y^2-\frac{2z}{v}y))dy \\
&= \frac{\exp(-\frac{z^2}{v})}{\sqrt{2\pi v}\sqrt{2\pi w}}\int_{-\infty}^{\infty} \frac{1}{\sqrt{\frac{1}{w}+\frac{1}{v}}}\exp(-\frac{1}{2}((s^2-\frac{2z}{v}\frac{1}{\sqrt{\frac{1}{w}+\frac{1}{v}}}s))ds \\
&= \frac{\exp(-\frac{z^2}{v})}{\sqrt{2\pi(v+w)}\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(-\frac{1}{2}((s-\frac{z}{v\sqrt{\frac{1}{z^2w}+\frac{1}{v}}})^2-(\frac{z}{v\sqrt{\frac{1}{w}+\frac{1}{v}}})^2))ds \\
&= \frac{1}{\sqrt{2\pi(v+w)}}\exp(\frac{-z^2}{2v})\exp(\frac{z^2}{2v^2(\frac{1}{v}+\frac{1}{w})})\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(-\frac{1}{2}(s-\frac{z}{v\sqrt{\frac{1}{v}+\frac{1}{w}}})^2)ds \\
&= \frac{1}{\sqrt{2\pi(v+w)}}\exp(-\frac{z^2}{2(v+w)})
\end{aligned}$$

where $-\infty < z < \infty$. Now, if $X \sim N(a, v)$ and $Y \sim N(b, w)$, then $X - a \sim N(0, v)$ and $Y - b \sim N(0, w)$. Hence, applying above result, we have that $Z - (a + b) \sim N(0, v + w)$, or equivalently , $Z \sim N(a + b, v + w)$. □

Chapter 10

Weak convergence

Exercise 10.3.2. Let $X, Y_1, Y_2, \cdots$ be independent random variables, with $P(Y_n = 1) = \frac{1}{n}$ and $P(Y_n = 0) = 1 - \frac{1}{n}$. Let $Z_n = X + Y_n$. Prove that $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(X)$, i.e. that the law of Z_n converges weakly to the law of X.

Solution. Given $\epsilon > 0$. Then, $\{|Z_n - X| \geq \epsilon\} = \phi \text{if}(\epsilon > 1), \{1\} \text{ if}(0 < \epsilon \leq 1)$, implying $0 \leq P(|Z_n - X| \geq \epsilon) \leq \frac{1}{n}$ for all $n \in \mathbb{N}$. Accordingly, $\lim_n P(|Z_n - X| \geq \epsilon) = 0$. Hence, $\lim_n Z_n = X$ in probability, and consequently, by Proposition 10.2.1., $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(X)$. $\square$

Exercise 10.3.4. Prove that weak limits, if they exist, are unique. That is, if $\mu, \nu, \mu_1, \mu_2, \cdots$ are probability measures, and $\mu_n \Rightarrow \mu$ and also $\mu_n \Rightarrow \nu$, then $\mu = \nu$.

Solution. Put $F_n(x) = \mu_n((-\infty, x])$, $F(x) = \mu((-\infty, x])$ and $G(x) = \nu((-\infty, x])$ for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$. Then, by Theorem 10.1.1., and Exercise 6.3.7. it follows that $\lim_n F_n(x) = F(x)$ except of D_F and $\lim_n F_n(x) = G(x)$ except of D_G. Since F and G are increasing, the sets D_F and D_G are countable, and so is $D_F \cup D_G$. Thus, $F = G$ except at most on $D_F \cup D_G$. On the other hand, F and G are right continuous everywhere and , in particular, on $D_F \cup D_G$, implying $F = G$ on $D_F \cup D_G$

(In fact, let $x \in D_F \cup D_G$ and $\{x_n\}_{n=1}^{\infty}$ be a sequence of $(D_F \cup D_G)^c$ such that $x_n \searrow x$. Then, by the recent result it follows that :

$$F(x) = \lim_n F(x_n) = \lim_n G(x_n) = G(x)).$$

Accordingly, $F = G$ on $(-\infty, \infty)$. Eventually, by Proposition 6.0.2., $\mu = \nu$. $\square$

Exercise 10.3.6. Let $A_1, a_2, \cdots$ be any sequence of non-negative real numbers with $\sum_i a_i = 1$. Define the discrete measure μ by $\mu(.) = \sum_{i \in \mathbb{N}} a_i \delta_i(.)$, where $\delta_i(.)$ is a point mass at the positive integer i. Construct a sequence $\{\mu_n\}$ of probability measures , each having a density with respect to Lebesgue measure , such that $\mu_n \Rightarrow \mu$.

Solution. Define:

$$\mu_n(.) = \sum_{i \in \mathbb{N}} n a_i \lambda(. \cap (i - \frac{1}{n}, i])$$

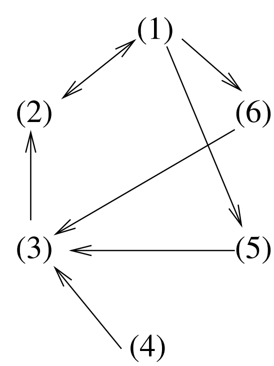
where λ dentes the Lebesgue measure. Then,

$$\mu_n((-\infty,\infty)) = 1,$$

for all $n \in \mathbb{N}$. Next, given $x \in \mathbb{R}$ and $\epsilon > 0$. Then, for any $n \in \mathbb{N}$ if $n \geq \frac{1}{1-(x-[x])}$, then $|\mu_n((-\infty,x]) - \mu((-\infty,x])| = 0 < \epsilon$. Accordingly, $\lim_n \mu_n((-\infty,x]) = \mu((-\infty,x])$. Eventually, by Theorem 10.1.1, $\mu_n \Rightarrow \mu$. $\square$

Exercise 10.3.8. Prove the following are equivalent:
(1) $\mu_n \Rightarrow \mu$
(2) $\int f d\mu_n \to \int f d\mu$, for all non-negative bounded continuous $f : \mathbb{R} \to \mathbb{R}$.
(3) $\int f d\mu_n \to \int f d\mu$, for all non-negative continuous $f : \mathbb{R} \to \mathbb{R}$with compact support, i.e., such that there are finite a and b with $f(x) = 0$ for all $x < a$ and all $x > b$.
(4) $\int f d\mu_n \to \int f d\mu$, for all continuous $f : \mathbb{R} \to \mathbb{R}$ with compact support.
(5) $\int f d\mu_n \to \int f d\mu$, for all non-negative continuous $f : \mathbb{R} \to \mathbb{R}$ which vanish at infinity, i.e., $\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(x) = 0$.
(6) $\int f d\mu_n \to \int f d\mu$, for all continuous $f : \mathbb{R} \to \mathbb{R}$ which vanish at infinity.

Solution. We prove the equivalence of the assertions according the following diagram:



(1) $\to$ (2) :
Any non-negative bounded continuous function is bounded continuous. Hence, by definition of weak convergence, (1) implies (2).

(1) $\to$ (4) :
Any continuous function f having compact support is bounded continuous, (since a continuous function attains its supremum on any compact set). Hence, by definition of weak convergence, (1) implies (4).

(1) $\to$ (5) :
Any positive continuous function which vanishes at $\pm\infty$ is bounded continuous(since there is a compact set $[-M, M]$ such that $|f| \leq 1$ outside of it, implying

$$|f| \leq \max(1, \sup_{[-M,M]} |f(x)|) < \infty).$$

Hence, by definition of weak convergence, (1) implies (5).

(1) $\to$ (6) :

Any continuous function which vanishes at $\pm\infty$ is bounded continuous(since there is a compact set $[-M, M]$ such that $|f| \leq 1$ outside of it, implying

$$|f| \leq \max(1, \sup_{[-M,M]} |f(x)|) < \infty).$$

Hence, by definition of weak convergence, (1) implies (5).

$(4) \to (3)$:
Any function satisfying (3) is an special case of (4). Thus, (4) implies (3).

$(5) \to (3)$:
Any function satisfying (3) is an special case of (5). Thus, (5) implies (3).

$(6) \to (3)$:
Any function satisfying (3) is an special case of (6). Thus, (6) implies (3).

$(2) \to (1)$:
Let f be a bounded continuous function. Then, $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$ are non-negative bounded continuous functions. Hence, by (2) it follows:

$$\lim_n \int f d\mu_n = \lim_n \int f^+ d\mu_n - \lim_n \int f^- d\mu_n = \int f^+ d\mu - \int f^+ d\mu = \int f d\mu.$$

$(3) \to (2)$:
Let f be a non-negative bounded continuous function. Put $f_M = f 1_{[-M,M]}$ for $M > 0$. We claim that:

$$\lim_n \int f_M d\mu_n = \int f_M d\mu \text{ for any } M > 0.(\star)$$

To prove the assertion, let $\epsilon > 0$. Then, there are non negative-compact support- continuous functions g_M and h_M with $h_M \leq f_M \leq g_M$ such that :

$$\begin{aligned} \int f_M d\mu &\leq \int g_M d\mu \leq \int f_M d\mu + \epsilon, \\ \int f_M d\mu &\geq \int h_M d\mu \geq \int f_M d\mu - \epsilon, \end{aligned}$$

implying:

$$\begin{aligned}
\int f_M d\mu - \epsilon &\leq \int h_M d\mu \\
&= \liminf_n \int h_M d\mu_n \\
&\leq \liminf_n \int f_M d\mu_n \\
&\leq \limsup_n \int f_M d\mu_n \\
&\leq \limsup_n \int g_M d\mu_n \\
&= \int g_M d\mu \\
&\leq \int f_M d\mu + \epsilon.
\end{aligned}$$

Now, since the $\epsilon > 0$ can be chosen arbitrary, the desired result follows. Next, let (2) does not hold. Therefore, there exist $\epsilon_0 > 0$ and an increasing sequence $\{n_k\}_{k=1}^{\infty}$ such that $|\int f d\mu_{n_k} - \int f d\mu| > \epsilon_0$ for all $k \in \mathbb{N}$. On the other hand, using $(\star)$ there are large enough $M_0 > 0$ and $k_0 \in \mathbb{N}$ such that :

$$\begin{aligned}
|\int f d\mu_{n_{k_0}} - \int f_{M_0} d\mu_{n_{k_0}}| &< \frac{\epsilon_0}{4}, \\
|\int f_{M_0} d\mu_{n_{k_0}} - \int f_{M_0} d\mu| &< \frac{\epsilon_0}{4}, \\
|\int f_{M_0} d\mu - \int f d\mu| &< \frac{\epsilon_0}{4},
\end{aligned}$$

Eventually, an application of triangular inequality shows that:

$$\begin{aligned}
\epsilon_0 < |\int f d\mu_{n_{k_0}} - \int f d\mu| &\leq |\int f d\mu_{n_{k_0}} - \int f_{M_0} d\mu_{n_{k_0}}| \\
&+ |\int f_{M_0} d\mu_{n_{k_0}} - \int f_{M_0} d\mu| \\
&+ |\int f_{M_0} d\mu - \int f d\mu| \\
&< \frac{3\epsilon_0}{4},
\end{aligned}$$

which is a contradiction. □

Exercise 10.3.10. Let $f : [0,1] \to (0,\infty)$ be a continuous function such that $\int_0^1 f d\lambda = 1$ (where λ is the Lebesgue measure on [0,1]). Define probability measures μ and $\{\mu_n\}$ by $\mu(A) = \int_0^1 f 1_A d\lambda$ and

$$\mu_n(A) = \sum_{i=1}^{n} f(i/n) 1_A(i/n) / \sum_{i=1}^{n} f(i/n).$$

(a) Prove that $\mu_n \Rightarrow \mu$.
(b) Explicitly, construct random variables Y and $\{Y_n\}$ so that $\mathcal{L}(Y) = \mu$, $\mathcal{L}(Y_n) = \mu_n$, and $Y_n \to Y$ with probability 1.

Solution.(a) Let A be a measurable set with $\mu(\partial A) = 0$. Then:

$$\begin{aligned}
\lim_n \mu_n(A) &= \lim_n \frac{\sum_{i=1}^n f(i/n) 1_A(i/n)}{\sum_{i=1}^n f(i/n)} \\
&= \lim_n \frac{\sum_{i=1}^n f(i/n) 1_A(i/n) 1/n}{\sum_{i=1}^n f(i/n) 1/n} \\
&= \frac{\lim_n \sum_{i=1}^n f(i/n) 1_A(i/n) 1/n}{\lim_n \sum_{i=1}^n f(i/n) 1/n} \\
&= \frac{\int_0^1 f 1_A d\lambda}{\int_0^1 f d\lambda} \\
&= \int_0^1 f 1_A d\lambda \\
&= \mu(A)
\end{aligned}$$

Consequently, by Theorem 10.1.1, $\mu_n \Rightarrow \mu$.

(b) Let $(\Omega, \mathcal{F}, P)$ be Lebesgue measure on $[0, 1]$. Put

$$F_n(x) = \mu_n((-\infty, x]) = \left(\frac{\sum_{i=1}^{[nx]} f(\frac{i}{n})}{\sum_{i=1}^n f(\frac{i}{n})}\right) 1_{[0,1]}(x) + 1_{(1,\infty)}(x)$$

and

$$F(x) = \mu((-\infty, x]) = \int_0^x f(t) dt 1_{[0,1]}(x) + 1_{(1,\infty)}(x).$$

Then, consider :

$$\begin{aligned}
Y_n(w) &= \inf\{x : F_n(x) \geq w\} \\
&= \inf\{x \in [0, 1] : \sum_{i=1}^{[nx]} f(\frac{i}{n}) \geq \sum_{i=1}^n f(\frac{i}{n}) w\},
\end{aligned}$$

and

$$\begin{aligned}
Y(w) &= \inf\{x : F(x) \geq w\} \\
&= \inf\{x \in [0, 1] : \int_0^x f(t) dt \geq w\},
\end{aligned}$$

where $0 \leq w \leq 1$. Now, by the proof of the Theorem 10.1.1, it follows that $Y_n \to Y$ with probability 1. $\square$

Chapter 11

Characteristic functions

Exercise 11.5.2. Let $\mu_n = \delta_{n \bmod 3}$ be a point mass at n mod 3.(Thus, $\mu_1 = \delta_1$, $\mu_2 = \delta_2$, $\mu_3 = \delta_0$, $\mu_4 = \delta_1$, $\mu_5 = \delta_2$, $\mu_6 = \delta_0$, etc.)
(a) Is $\{\mu_n\}$ tight?
(b) Does there exist a Borel probability measure μ, such that $\mu_n \Rightarrow \mu$? (If so, then specify μ.)
(c) Does there exist a subsequence $\{\mu_{n_k}\}$, and a Borel probability measure μ , such that $\mu_{n_k} \Rightarrow \mu$? (If so, then specify $\{\mu_{n_k}\}$ and μ.)
(d) Relate parts (b) and (c) to theorems from this section.

Solution. (a) Yes. Take $[a, b] = [0, 3]$, then, for all $\epsilon > 0$ and for all $n \in \mathbb{N}$:

$$\mu_n([0,3]) = \delta_n([0,3])_{\text{mod } 3} = 1_{[0,3]}(n)|_{n=0,1,2} = 1 > 1 - \epsilon.$$

(b) No. Assume, there exists such a distribution μ. Hence, by Theorem 10.1.1, for any $x \in \mathbb{R}$ with $\mu(\{x\}) = 0$, it follows $\lim_n \mu_n((-\infty, x]) = \mu((-\infty, x])$. On the other hand, pick $x \in (0, 1)$ such that $\mu(\{x\}) = 0,$. Then, $\mu_n((-\infty, x]) = 1 \mathrm{if} 3|n, 0 \mathrm{etc.}$ for all $n \in \mathbb{N}$. Therefore, $\lim_n \mu_n((-\infty, x])$ does not exist, a contradiction.

(c) Yes. Put $n_k = 3k + 1$ for $k = 0, 1, 2, \ldots$. Then, $\mu_{n_k} = \mu_1 = \delta_1$ for $k = 0, 1, 2, \ldots$, implying $\mu_{n_k} \Rightarrow \delta_1$.

(d) The sequence $\{\mu_n\}_{n=1}^{\infty}$ is a tight sequence of probability measures, satisfying assumptions of the Theorem 11.1.10, and according to that theorem, there exists a subsequence $\{\mu_{n_k}\}_{k=1}^{\infty}$ (here, $n_k = 3k + 1$ for $k \in \mathbb{N}$) such that for some probability measure μ (here, $\mu = \delta_1$), $\mu_{n_k} \Rightarrow \mu$. $\square$

Exercise 11.5.4. Let $\mu_{2n} = \delta_0$, and let $\mu_{2n+1} = \delta_n$, for $n = 1, 2, \cdots$.
(a) Does there exist a Borel probability measure μ such that $\mu_n \Rightarrow \mu$?
(b) Suppose for some subsequence $\{\mu_{n_k}\}_{k=1}^{\infty}$ and some Borel probability measure ν, we have $\mu_{n_k} \Rightarrow \nu$. What must ν must be?
(c) Relate parts (a) and (b) to Corollary 11.1.11. Why is there no contradiction?

Solution. (a) No. Assume, there exists such a distribution μ. Hence, by Theorem 10.1.1, for any $x \in \mathbb{R}$ with $\mu(\{x\}) = 0$, it follows $\lim_n \mu_n((-\infty, x]) = \mu((-\infty, x])$. On the other hand, pick $x \in (0, 1)$ such that $\mu(\{x\}) = 0$.

Then, $\mu_n((-\infty,x]) = 1 \text{if} 2|n, 0 \text{etc.}$ for all $n \in \mathbb{N}$. Therefore, $\lim_n \mu_n((-\infty,x])$ does not exist, a contradiction.

(b) Since any subsequence $\{\mu_{n_k}\}_{k=1}^{\infty}$ including an infinite subsequence of $\{\mu_{2k+1}\}_{k=1}^{\infty}$ is not weakly convergence, we must consider $\{\mu_{n_k}\}_{k=1}^{\infty}$ as a subsequence of $\{\mu_{2k}\}_{k=1}^{\infty}$. In this case, $\mu_{n_k} = \nu = \delta_0$ for all $k \in \mathbb{N}$, implying $\mu_{n_k} \Rightarrow \nu$.

(c) Since $\{\mu_n\}_{n=1}^{\infty}$ is not tight (Take $\epsilon = \frac{1}{2}$. Then, for any $[a,b] \subseteq \mathbb{R}$ and any $n \in \mathbb{N}$, if $n \geq \lfloor b \rfloor + 1$ then $\mu_{2n+1}([a,b]) = 0 < 1 - \frac{1}{2}$.), the hypothesis of Corollary 11.1.11 is violated. Consequently, there is no contradiction to the assertion of that Corollary.□

Exercise 11.5.6. Suppose $\mu_n \Rightarrow \mu$. Prove or disprove that $\{\mu_n\}$ must be tight.

Solution. Method(1):
Given $\epsilon > 0$. Since $\mu(\mathbb{R}) = 1$, there is a closed interval $[a_0, b_0]$ such that:

$$1 - \frac{\epsilon}{2} \leq \mu([a_0, b_0]). \qquad (\star)$$

Next, by Theorem 10.1.1(2), there is a positive integer N such that:

$$|\mu_n([a_0,b_0]) - \mu([a_0,b_0])| \leq \frac{\epsilon}{2} \qquad n = N+1, \cdots,$$

implying:

$$\mu([a_0,b_0]) - \frac{\epsilon}{2} \leq \mu_n([a_0,b_0]) \qquad n = N+1, \cdots. \qquad (\star\star)$$

Combining $(\star)$ and $(\star\star)$ yields:

$$1 - \epsilon \leq \mu_n([a_0,b_0]) \qquad n = N+1, \cdots. \qquad (\star\star\star)$$

Next, for each $1 \leq n \leq N$, there is a closed interval $[a_n, b_n]$ such that:

$$1 - \epsilon \leq \mu_n([a_n,b_n]) \qquad n = 1, \cdots, N. \qquad (\star\star\star\star)$$

Define:

$$a = \min_{0 \leq n \leq N}(a_n) \qquad b = \max_{0 \leq n \leq N}(b_n).$$

Then, by $[a_n,b_n] \subseteq [a,b]$ for all $0 \leq n \leq N$, the inequality $(\star\star\star)$ and the inequality $(\star\star\star\star)$ we have that :

$$1 - \epsilon \leq \mu_n([a,b]) \qquad n = 1, 2, \cdots.$$

Method(2):
Let $\phi, \phi_1, \phi_2, \cdots$, be corresponding characteristic functions to the probability measures $\mu, \mu_1, \mu_2, \cdots$. Then, by Theorem 11.1.14(the continuity theorem) $\mu_n \Rightarrow \mu$ implies, $\lim_n \phi_n(t) = \phi(t)$ for all $t \in \mathbb{R}$. On the other hand, ϕ is continuous on the $\mathbb{R}$. Thus, according to Lemma 11.1.13, $\{\mu_n\}$ is tight. □

Exercise 11.5.8. Use characteristic functions to provide an alternative solution of Exercise 10.3.2.

Solution. Since X and Y_n are independent, it follows:

$$\phi_{Z_n}(t) = \phi_X(t)\phi_{Y_n}(t) = \phi_X(t)(\exp(it).\frac{1}{n} + (1 - \frac{1}{n}))\text{for all}(n \in \mathbb{N}).$$

Therefore, $\lim_n \phi_{Z_n}(t) = \phi_X(t)$ for all $t \in \mathbb{R}$. Hence, by Theorem 11.1.14 (the continuity theorem), $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(X)$. $\square$

Exercise 11.5.10. Use characteristic functions to provide an alternative solution of Exercise 10.3.4.

Solution. Let $\mu_n \Rightarrow \mu$ and $\mu_n \Rightarrow \nu$. Then, by Theorem 11.1.14, $\lim_n \phi_n(t) = \phi_\mu(t)$ and $\lim_n \phi_n(t) = \phi_\nu(t)$ for all $t \in \mathbb{R}$. Hence, $\phi_\mu(t) = \phi_\nu(t)$ for all $t \in \mathbb{R}$. Eventually, by Corollary 11.1.7., $\mu = \nu$. $\square$

Exercise 11.5.12. Suppose that for $n \in \mathbb{N}$, we have $P(X_n = 5) = \frac{1}{n}$ and $P(X_n = 6) = 1 - \frac{1}{n}$.
(a) Compute the characteristic function $\phi_{X_n}(t)$, for all $n \in \mathbb{N}$ and all $t \in \mathbb{R}$.
(b) Compute $\lim_{n\to\infty} \phi_{X_n}(t)$.
(c) Specify a distribution μ such that $\lim_{n\to\infty} \phi_{X_n}(t) = \int \exp(itx)\mu(dx)$ for all $t \in \mathbb{R}$.
(d) Determine (with explanation) whether or not $\mathcal{L}(X_n) \Rightarrow \mu$.

Solution. (a)
$\phi_{X_n}(t) = E(\exp(itX_n)) = \frac{\exp(5it)}{n} + \frac{(n-1)\exp(6it)}{n}$ $t \in \mathbb{R}$.
(b)
$\lim_{n\to\infty} \phi_{X_n}(t) = \exp(6it)$ $t \in \mathbb{R}$.
(c)
$\mu = \delta_6$.
(d)
As a result of the Theorem 11.1.14 (the continuity theorem), $\mathcal{L}(X_n) \Rightarrow \delta_6$. $\square$

Exercise 11.5.14. Let $X_1, X_2, \cdots$ be i.i.d. with mean 4 and variance 9. Find values $C(n, x)$, for $n \in \mathbb{N}$ and $x \in \mathbb{R}$, such that as $n \to \infty$, $P(X_1 + X_2 + \cdots + X_n \leq C(n, x)) \approx \Phi(x)$.

Solution. As in Corollary 11.2.3:

$$P(\frac{S_n - nm}{\sqrt{nv}} \leq x) = \Phi(x) \qquad \text{as} n \to \infty$$

Thus :

$$\Phi(x) = P(\frac{S_n - 4n}{\sqrt{9n}} \leq \frac{C(n,x) - 4n}{\sqrt{9n}}) = \Phi(\frac{C(n,x) - 4n}{\sqrt{9n}}),$$

as$n \to \infty$, and injectivity of Φ implies :

$$x = \frac{C(n,x) - 4n}{\sqrt{9n}},$$

or :

$$C(n,x) = \sqrt{9n}x + 4n.$$

$\square$

Exercise 11.5.16. Let X be a random variable whose distribution $\mathcal{L}(X)$ is infinitely divisible. Let $a > 0$ and $b \in \mathbb{R}$, and set $Y = aX + b$, prove that $\mathcal{L}(Y)$ is infinitely divisible.

Solution. Given $n \in \mathbb{N}$. Since the distribution $\mathcal{L}(X)$ is infinitely divisible, there is a distribution ν_n such that if $X_k \sim \nu_n$, $1 \leq k \leq n$, are independent random variables, then $\sum_{k=1}^n X_k \sim \mathcal{L}(X)$ or, equivalently, there is a distribution ν_n such that if X_k , $1 \leq k \leq n$, with $F_{X_k}(x) = \nu_n((-\infty, x]), (x \in \mathbb{R})$ are independent, then $F_{\sum_{k=1}^n X_k}(x) = F_X(x), (x \in \mathbb{R})$. The latest assertion yields that there is a distribution ω_n defined by $\omega_n((-\infty, x]) = \nu_n((-\infty, \frac{x-b}{a}]), (x \in \mathbb{R})$such that if X_k , $1 \leq k \leq n$, with $F_{aX_k+b}(x) = \omega_n((-\infty, x]), (x \in \mathbb{R})$ are independent, then $F_{\sum_{k=1}^n (aX_k+b)}(x) = F_{aX+b}(x), (x \in \mathbb{R})$. Accordingly, the distribution $\mathcal{L}(Y)$ is infinitely divisible, as well.□

Exercise 11.5.18. Let $X, X_1, X_2, \cdots$ be random variables which are uniformly bounded , i.e. there is $M \in \mathbb{R}$ with $|X| \leq M$ and $|X_n| \leq M$ for all n. Prove that $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$ if and only if $E(X_n^k) \to E(X^k)$ for all $k \in \mathbb{N}$.

Solution. Let $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$. Then, by Theorem 10.1.1, $\lim_n \int f d(\mathcal{L}(X_n)) = \int f d(\mathcal{L}(X))$ for any bounded Borel measurable function f with $\mathcal{L}(X)(D_f) = 0$. Put $f_k(x) = x^k (k \in \mathbb{N})$. Then, since $|f_k| \leq M^k < \infty (k \in \mathbb{N})$, it follows that:

$$\lim_n E(X_n^k) = \lim_n \int x^k d(\mathcal{L}(X_n)) = \int x^k d(\mathcal{L}(X)) = E(X^k),$$

for all $k \in \mathbb{N}$.

Conversely, let $E(X_n^k) \to E(X^k)$ for all $k \in \mathbb{N}$. Since $|X| \leq M$,

$$|M_X(s)| = |E(e^{sX})| \leq \max(e^{-M}, e^M) < \infty$$

for all $|s| < 1$. Therefore, by Theorem 11.4.3, $\mathcal{L}(X)$ is determined by its moments. Now, the desired assertion follows from Theorem 11.4.1. □

Chapter 12

Decompositions of probability laws

Exercise 12.3.2. Let X and Y be discrete random variables (not necessarily independent), and let $Z = X + Y$. Prove that $\mathcal{L}(Z)$ is discrete.

Solution. Let F be an an uncountable set of positive numbers. Then, $\sum_{x\in F} x = +\infty$. Thus, the summation in the definition of Borel probability measure of discrete random variable is, in fact, taken over a countable set for which the probability of each of its elements is nonzero. Hence, in the equalities

$$\sum_{x\in\mathbb{R}} \mathcal{L}(X)(\{x\}) = \mathcal{L}(X)(\mathbb{R})$$

and

$$\sum_{y\in\mathbb{R}} \mathcal{L}(Y)(\{y\}) = \mathcal{L}(Y)(\mathbb{R}),$$

the sets $A_X = \{x \in \mathbb{R} : \mathcal{L}(X)(\{x\}) > 0\}$ and $B_Y = \{y \in \mathbb{R} : \mathcal{L}(Y)(\{y\}) > 0\}$ are countable. On the other hand, for any $z \in \mathbb{R}$ we have $\mathcal{L}(X+Y)(\{z\}) \geq P(X = x, Y = z - x) > 0 \text{for some} x \in \mathbb{R}$, and the corresponding ordered pair $(x, z - x)$ in the recent result is an element of the countable set $A_X \times B_Y$. Hence, the set $C_Z = \{z \in \mathbb{R} : \mathcal{L}(X+Y)(\{z\}) > 0\}$ is countable. Thus:

$$\begin{aligned}
\sum_{z\in\mathbb{R}} \mathcal{L}(X+Y)(\{z\}) &= \sum_{z\in C_Z} \mathcal{L}(X+Y)(\{z\}) \\
&= \mathcal{L}(X+Y)(C_Z) + 0 \\
&= \mathcal{L}(X+Y)(C_Z) + \mathcal{L}(X+Y)(C_Z^c) \\
&= \mathcal{L}(X+Y)(\mathbb{R}).
\end{aligned}$$

□

Exercise 12.3.4. Let X and Y be random variables, with $\mathcal{L}(Y)$ absolutely continuous, and let $Z = X + Y$.
(a) Assume X and Y are independent. Prove that $\mathcal{L}(Z)$ is absolutely continuous, regardless of the nature of $\mathcal{L}(X)$.
(b) Show that if X and Y are not independent, then $\mathcal{L}(Z)$ may fail to be absolutely continuous.

Solution. (a) Let for a Borel measurable set $A \subseteq \mathbb{R}$, $\lambda(A) = 0$. Hence, $\lambda(A - x) = 0$, for all $x \in \mathbb{R}$. Then, by Corollary 12.1.2.(Radon-Nikodym Theorem), $\mathcal{L}(Y)(A - x) = 0$ for all $x \in \mathbb{R}$. Now, by Theorem 9.4.5.:

$$\mathcal{L}(Z)(A) = \int \mathcal{L}(Y)(A - x)\mathcal{L}(X)dx = 0.$$

Eventually, by another application of the Corollary 12.1.2.(Radon-Nikodym Theorem), the desired result follows.

(b) First example:
Put $X = Z$ and $Y = -Z$, where $Z \sim N(0,1)$. Then, $X + Y = 0$, which is clearly discrete.

Second example:
Similar to Exercise 6.3.5., let $X \sim N(0,1)$, $Y = XU$ where X and U are independent and $P(U = 1) = P(U = -1) = \frac{1}{2}$. Then, X and Y are dependent. Put $Z = X + Y$. Then, for the Borel measurable set $A = \{0\}$, $\lambda(A) = 0$. But, $\mathcal{L}(Z)(A) = \frac{1}{2} \neq 0$ (Check!).$\square$

Exercise 12.3.6. Let $A, B, Z_1, Z_2, \cdots$ be i.i.d., each equal to $+1$ with probability $2/3$, or equal to 0 with probability $1/3$. Let $Y = \sum_{i=1}^{\infty} Z_i 2^{-i}$ as at the beginning of this section (so $\nu = \mathcal{L}(Y)$ is singular continuous), and let $W \sim N(0,1)$. Finally, let $X = A(BY + (1-B)W)$, and set $\mu = \mathcal{L}(X)$. Find a discrete measure μ_{disc}, an absolutely continuous measure μ_{ac}, and a singular continuous measure μ_s, such that $\mu = \mu_{disc} + \mu_{ac} + \mu_s$.

Solution. Using conditional probability, for any Borel measurable set $U \subseteq \mathbb{R}$, it follows:

$$\begin{aligned}
\mathcal{L}(X)(U) &= P(X \in U) \\
&= P(X \in U | A = 0)P(A = 0) \\
&+ P(X \in U | A = 1)P(A = 1) \\
&= \frac{1}{3}P(0 \in U) \\
&+ \frac{2}{3}P(BY + (1-B)W \in U) \\
&= \frac{1}{3}\delta_0(U) \\
&+ \frac{2}{3}(P(BY + (1-B)W \in U | B = 0)P(B = 0) \\
&+ P(BY + (1-B)W \in U | B = 1)P(B = 1)) \\
&= \frac{1}{3}\delta_0(U) + \frac{2}{3}(\frac{1}{3}P(W \in U) + \frac{2}{3}P(Y \in U)) \\
&= \frac{1}{3}\delta_0(U) + \frac{2}{9}\mathcal{L}(W)(U) + \frac{4}{9}\mathcal{L}(Y)(U).
\end{aligned}$$

Accordingly, $\mathcal{L}(X) = \frac{1}{3}\delta_0 + \frac{2}{9}\mathcal{L}(W) + \frac{4}{9}\mathcal{L}(Y)$. $\square$

Exercise 12.3.8. Let μ and ν be probability measures with $\mu \ll \nu$ and $\nu \ll \mu$. (This is sometimes written as $\mu \equiv \nu$.) Prove that $\frac{d\mu}{d\nu} > 0$ with μ-probability 1, and in fact $\frac{d\nu}{d\mu} = \frac{1}{\frac{d\mu}{d\nu}}$.

Solution. By Exercise 12.3.7., $\frac{d\mu}{d\rho} = \frac{d\mu}{d\nu}\frac{d\nu}{d\rho}$ with ρ-probability 1. We claim that :

$$\frac{d\mu}{d\nu} = \frac{\frac{d\mu}{d\rho}}{\frac{d\nu}{d\rho}} \qquad \nu - \text{probability1.}(\star)$$

To prove the above equality we notice that :

$$\nu(\{w : \frac{d\nu}{d\rho} = 0\}) = \int_{\{w: \frac{d\nu}{d\rho}=0\}} \frac{d\nu}{d\rho} d\rho = 0,$$

consequently, $\frac{d\nu}{d\rho} > 0, \nu -$ probability-1 . Hence, $(\star)$ holds $\nu -$ probability-1. Of course on the set $\{w : \frac{d\nu}{d\rho} = 0\}$, the right hand side of $(\star)$ can be defined as zero. Now, let $\rho = \mu$. Then, $\frac{d\nu}{d\mu} > 0, \nu -$ probability-1 and $\frac{d\mu}{d\nu} = \frac{1}{\frac{d\nu}{d\mu}}$. Eventually, the final assertion follows by symmetric relationship of μ and ν. □

Exercise 12.3.10. Let μ and ν be absolutely continuous probability measures on $\mathbb{R}$ (with the Borel σ algebra), with $\phi = \mu - \nu$. Write down an explicit Hahn decomposition $\mathbb{R} = A^+ \cup A^-$ for ϕ.

Solution. The solution does not answer the question as it does not give A+ and A-.
Since μ is an absolutely continuous probability measure on $\mathbb{R}$, it is countably additive. Indeed, let $A_1, A_2, \ldots$ be disjoint sets in the Borel σ-algebra,

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \int 1_{\bigcup_{n=1}^{\infty} A_n} \frac{d\mu}{d\lambda} d\lambda \\ &= \int \sum_{n=1}^{\infty} 1_{A_n} \frac{d\mu}{d\lambda} d\lambda \\ &= \sum_{n=1}^{\infty} \int 1_{A_n} \frac{d\mu}{d\lambda} d\lambda \\ &= \sum_{n=1}^{\infty} \mu(A_n)\end{aligned}$$

Similarly, ν is also countably additive. It follows from linearity of measures that $\phi = \mu - \nu$ is countably additive. By Remark (3) on Page 145, Lemma 12.1.4 applies to any countably additive mapping from $\mathcal{F}$ to $\mathbb{R}$. Hence, constructing A^+ as in the proof of Lemma 12.1.4 gives desired A^+. Set $\alpha = \sup\{\phi(A) : A \in \mathcal{F}\}$. Then, choose subsets $A_1, A_2, \ldots \in \mathcal{F}$ such that $\phi(A_n) \to \alpha$. Let $A = \bigcup A_i$ and

$$\mathcal{G}_n = \left\{ \bigcap_{k=1}^{n} A'_k, \text{ each } A'_k = A_k \text{ or } A \backslash A_k \right\}.$$

Define $C_n = \bigcup_{\substack{S \in \mathcal{G}_n \\ \phi(S) \geq 0}} S$. $A^+ = \limsup C_n$, $A^- = \Omega \backslash A^+$ gives the Hahn decomposition for ϕ.□

Chapter 13

Conditional probability and expectation

Exercise 13.4.2. Let $\mathcal{G}$ be a sub-σ-algebra, and let A be any event. Define the random variable X to be the indicator function 1_A. Prove that $E(X|\mathcal{G}) = P(A|\mathcal{G})$ with probability 1.

Solution. By definition, both of $E(1_A|\mathcal{G})$ and $P(A|\mathcal{G})$ are $\mathcal{G}$ measurable random variables. Next, by equations (13.2.1) and (13.2.2):

$$\begin{aligned} E(P(A|\mathcal{G})1_G) &= P(A \cap G) \\ &= E(1_{A\cap G}) \\ &= E(1_A 1_G) \\ &= E(E(1_A|\mathcal{G})1_G), \end{aligned}$$

for any $G \in \mathcal{G}$. Consequently, $E(1_A|\mathcal{G}) = P(A|\mathcal{G})$ with probability 1. □

Exercise 13.4.4. Let X and Y be random variables with joint distribution given by $\mathcal{L}(X,Y) = dP = f(x,y)\lambda_2(dx,dy)$, where λ_2 is two dimensional Lebesgue measure, and $f : \mathbb{R}^2 \to \mathbb{R}$ is a non-negative Borel measurable function with $\int_{\mathbb{R}^2} f d\lambda_2 = 1$. Show that we can take $P(Y \in B|X) = \int_B g_X(y)\lambda(dy)$ and $E(Y|X) = \int_{\mathbb{R}} y g_X(y)\lambda(dy)$, where the function $g_x : \mathbb{R} \to \mathbb{R}$ is defined by $g_x(y) = \frac{f(x,y)}{\int_{\mathbb{R}} f(x,t)\lambda(dt)}$ whenever $\int_{\mathbb{R}} f(x,t)\lambda(dt)$ is positive and finite, otherwise (say) $g_x(y) = 0$.

Solution. First, let $\widehat{Y_B(x)} = \int_B g_X(y)\lambda(dy)$. Then, $\widehat{Y_B(x)}$ is a $\sigma(X)$-measurable random variable since $\int_B g_x(y)\lambda(dy)$ is a Borel function of $x \in \mathbb{R}$. $f_X(x) = \int_{\mathbb{R}} f(x,y)\lambda(dy)$ and is Borel measurable as $f(x,y)$ is Borel measurable. For $A = (Y \in B)$ in Definition 13.1.4, an application of the Fubini's Theorem in Real Analysis yields:

$$\begin{aligned} E\left(\widehat{Y_B(x)}1_{X\in S}\right) &= E\left(\int_B g_X(y)\lambda(dy)1_{X\in S}\right) \\ &= \int_S \int_B g_X(y) f_X(x)\lambda(dy)\lambda(dx) \\ &= \int\int_{(S\times B)} f(x,y)\lambda_2(dx,dy) \\ &= P((Y \in B) \cap \{X \in S\}), \end{aligned}$$

for all Borel $S \subseteq \mathbb{R}$, proving that $P(Y \in B|X) = \int_B g_X(y)\lambda(dy)$ with probability 1.
Second, by considering $E(1_B|X) = P(Y \in B|X)$, the second equation follows from the first equation for the special case $Y = 1_B$. Now, by usual linearity and monotone convergence arguments the second equality follows for general Y. $\square$

Exercise 13.4.6. Let $\mathcal{G}$ be a sub-σ-algebra, and let X and Y be two independent random variables. Prove by example that we might have:

$$E(XY|\mathcal{G}) \neq E(X|\mathcal{G})E(Y|\mathcal{G}).$$

Solution. Let $(\Omega, \mathcal{F}, P)$ be the Lebesgue measure on $[0,1]$. Consider the two independent events $A = [0, \frac{3}{9}],\ B = [\frac{2}{9}, \frac{5}{9}]$. Then, $P(A) = P(B) = \frac{1}{3}, P(A \cap B) = \frac{1}{9},\ \ P(A \cup B) = \frac{5}{9}$ and for $C = (A \cup B)^c$ we have: $P(C) = \frac{4}{9}$.

Next, consider the sub-$\sigma-$algebra $\mathcal{G} = \{\phi, C, C^c, \Omega\}$. Then, for any random variable X:

$$\begin{aligned} E(X|\mathcal{G}) &: \Omega \to \mathbb{R} \\ E(X|\mathcal{G})(w) &= E(X|C).1_C(w) + E(X|C^c).1_{C^c}(w). \quad (*) \end{aligned}$$

In particular, for $X = 1_U$, we have:

$$\begin{aligned} P(U|\mathcal{G}) &: \Omega \to \mathbb{R} \\ P(U|\mathcal{G})(w) &= P(U|C).1_C(w) + P(U|C^c).1_{C^c}(w). \quad (**) \end{aligned}$$

Accordingly, for $X = 1_A, Y = 1_B$ $(X.Y = 1_{A\cap B})$, by three applications of $(**)$ we have:

$$\begin{aligned} E(XY|\mathcal{G})(w) &= P(A \cap B|C).1_C(w) + P(A \cap B|C^c).1_{C^c}(w) \\ &= 0 + \frac{P(A \cap B)}{P(A \cup B)}.1_{C^c}(w) = \frac{1/9}{5/9}.1_{C^c}(w) \\ &= \frac{1}{5}1_{C^c}(w), \\ E(X|\mathcal{G})(w) &= P(A|C).1_C(w) + P(A|C^c).1_{C^c}(w) \\ &= 0 + \frac{P(A)}{P(A \cup B)}.1_{C^c}(w) = \frac{1/3}{5/9}.1_{C^c}(w) \\ &= \frac{3}{5}1_{C^c}(w), \\ E(Y|\mathcal{G})(w) &= P(B|C).1_C(w) + P(B|C^c).1_{C^c}(w) \\ &= 0 + \frac{P(B)}{P(A \cup B)}.1_{C^c}(w) = \frac{1/3}{5/9}.1_{C^c}(w) \\ &= \frac{3}{5}1_{C^c}(w), \end{aligned}$$

yielding:

$$E(XY|\mathcal{G}) = \frac{1}{5}1_{C^c} \neq \frac{9}{25}1_{C^c} = (\frac{3}{5}1_{C^c}).(\frac{3}{5}1_{C^c}) = E(X|\mathcal{G}).E(Y|\mathcal{G}).$$

$\square$

Exercise 13.4.8. Suppose Y is $\mathcal{G}$-measurable. Prove that $Var(Y|\mathcal{G}) = 0$.

Solution. Since Y and $E(Y|\mathcal{G})$ are $\mathcal{G}$-measurable, it follows that $Y - E(Y|\mathcal{G})$ is $\mathcal{G}$-measurable, too. Now, an application of Proposition 13.2.6 yields:

$$\begin{aligned} Var(Y|\mathcal{G}) &= E((Y - E(Y|\mathcal{G}))^2|\mathcal{G}) \\ &= (Y - E(Y|\mathcal{G}))E(Y - E(Y|\mathcal{G})|\mathcal{G}) \\ &= (Y - E(Y|\mathcal{G}))(E(Y|\mathcal{G}) - E(Y|\mathcal{G})) \\ &= 0. \end{aligned}$$

□

Exercise 13.4.10. Give an example of jointly defined random variables which are not independent, but such that $E(Y|X) = E(Y)$ w.p. 1.

Solution. Let X, Y be jointly defined by:

$$P(X = 1, Y = 1) = P(X = 1, Y = -1) = P(X = -1, Y = 2) = P(X = -1, Y = -2) = \frac{1}{4}.$$

Since:

$$\begin{aligned} P(X = 1) &= \sum_{j=-\infty}^{\infty} P(X = 1, Y = j) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \\ P(Y = 1) &= \sum_{i=-\infty}^{\infty} P(X = i, Y = -1) = \frac{1}{4}, \end{aligned}$$

it follows that:

$$P(X = 1, Y = 1) = \frac{1}{4} \neq \frac{1}{8} = P(X = 1) * P(Y = 1),$$

showing that X, Y are dependent.

Next, by definition:

$$E(Y) = \sum_{x=-\infty, y=-\infty}^{+\infty, +\infty} y.P(X = x, Y = y) = \frac{-2 - 1 + 1 + 2}{4} = 0.$$

On the other hand, by two applications of Exercise 13.4.3(a) it follows that:

$$\begin{aligned} E(Y|X = -1) &= \frac{\sum_{y=-\infty}^{+\infty} y.P(X = -1, Y = y)}{\sum_{z=-\infty}^{+\infty} P(X = -1, Y = z)} = \frac{(2 - 2)/4}{(1 + 1)/4} = 0, \\ E(Y|X = 1) &= \frac{\sum_{y=-\infty}^{+\infty} y.P(X = 1, Y = y)}{\sum_{z=-\infty}^{+\infty} P(X = 1, Y = z)} = \frac{(1 - 1)/4}{(1 + 1)/4} = 0. \end{aligned}$$

Thus, $E(Y|X = x) = 0 (a.e)$. Consequently, $E(Y) = 0 = E(Y|X)$ $(w.p.1)$.

□

Exercise 13.4.12. Let $\{Z_n\}$ be independent, each with finite mean. Let $X_0 = a$, and $X_n = a + Z_1 + ... + Z_n$ for $n \geq 1$. Prove that

$$E(X_{n+1}|X_0, X_1, ..., X_n) = X_n + E(Z_{n+1}).$$

Solution. Let $\mathcal{G} = \sigma(X_0, X_1, ..., X_n)$. Then, X_n and $E(Z_{n+1})$ are $\mathcal{G}$ measurable and so is $X_n + E(Z_{n+1})$.
Claim. Z_{n+1} and 1_G are independent.
First, we will prove X_n and Z_k are independent for $k \leq n + 1$ by induction:

- Base case: <u>$n = 1$:</u> Since $Z_0 = a$, Z_1 and Z_k are independent for $k \geq 2$, we have $X_1 = a + Z_1$ and Z_k are independent by Exercise 3.6.6 for $k \geq 2$.
  <u>$n = 2$:</u> Since $X_1 = a + Z_1$, Z_2, Z_k are independent for $k \geq 3$, we have $X_2 = X_1 + Z_2 = a + Z_1 + Z_2$ and Z_k are independent by Exercise 3.6.6. for $k \geq 3$.
- Inductive step: Assume X_{n-1} is independent of Z_k for $k \geq n$. Since X_{n-1}, Z_n, and Z_k are independent for $k \geq n+1$, we have $X_n = X_{n-1} + Z_n = a + Z_1 + \cdots + Z_{n-1} + Z_n$ and Z_k are independent by Exercise 3.6.6 for $k \geq n+1$.

It follows that $X_1, \ldots, X_n$ is independent of Z_{n+1}. Therefore, 1_G is independent of Z_{n+1}.

Consequently:

$$\begin{aligned} E((X_n + E(Z_{n+1}))1_G) &= E(X_n 1_G) + E(E(Z_{n+1})1_G) \\ &= E(X_n 1_G) + E(Z_{n+1}1_G) \\ &= E((X_n + Z_{n+1})1_G) \\ &= E(X_{n+1}1_G), \end{aligned}$$

for all $G \in \mathcal{G}$. Thus, $E(X_{n+1}|X_0, X_1, ..., X_n) = X_n + E(Z_{n+1})$ w.p.1. $\square$

Chapter 14

Martingales

Exercise 14.4.2. Let $\{X_n\}$ be a submartingale, and let $a \in \mathbb{R}$. Let $Y_n = \max(X_n, a)$. Prove that $\{Y_n\}$ is also a submartingale.

Solution. The solution states

$$\mathbf{E}(Y_{n+1}|\sigma(Y_0, Y_1, \ldots, Y_n)) = \mathbf{E}(\max(X_{n+1}, a)|\sigma(X_0, X_1, \ldots, X_n)).$$

I wonder it this is wrong, because for the Y in this exercise

$$\sigma(Y_0, Y_1, \ldots, Y_n) \neq \sigma(X_0, X_1, \ldots, X_n).$$

Since $\sigma(Y_0, Y_1, \ldots, Y_n) \subset \sigma(X_0, X_1, \ldots, X_n)$, it follows from Proposition 13.2.7 that

$$\mathbf{E}(Y_{n+1}|\sigma(Y_0, Y_1, \ldots, Y_n)) = \mathbf{E}(\mathbf{E}(Y_{n+1}|\sigma(X_0, X_1, \ldots, X_n))|\sigma(Y_0, Y_1, \ldots, Y_n)),$$

with probability 1. Additionally, notice that

$$\begin{aligned}
\mathbf{E}(Y_{n+1}|\sigma(X_0, X_1, \ldots, X_n)) &= \mathbf{E}(\max(X_{n+1}, a)|\sigma(X_0, X_1, \ldots, X_n)) \\
&\geq \max(\mathbf{E}(X_{n+1}|\sigma(X_0, X_1, \ldots, X_n), a) && \text{(Conditional Jensen's Inequality)} \\
&\geq \max(X_n, a) && \text{(X_n is a submartingale)} \\
&= Y_n
\end{aligned}$$

Therefore,

$$\mathbf{E}(\mathbf{E}(Y_{n+1}|\sigma(X_0, X_1, \ldots, X_n))|\sigma(Y_0, Y_1, \ldots, Y_n)) \geq \mathbf{E}(Y_n|\sigma(Y_0, Y_1, \ldots, Y_n)),$$

with probability 1 and then

$$\mathbf{E}(Y_{n+1}|\sigma(Y_0, Y_1, \ldots, Y_n)) \geq \mathbf{E}(Y_n|\sigma(Y_0, Y_1, \ldots, Y_n)) = Y_n,$$

with probability 1 □

Exercise 14.4.4. The conditional Jensen's inequality states that if ϕ is a convex function, then $E(\phi(X)|\mathcal{G}) \geq \phi(E(X|\mathcal{G}))$.
(a) Assuming this, prove that if $\{X_n\}$ is a submartingale, then so is $\{\phi(X_n)\}$ whenever ϕ is non-decreasing and convex with $E(|\phi(X_n)|) < \infty$, for all n.
(b) Show that the conclusion of the two previous exercises follow from part (a).

Solution. (a)

Same issue with Exercise 14.4.2

Claim. $\sigma(\phi(X_0), \ldots, \phi(X_n)) \subset \sigma(X_0, \ldots, X_n)$.
Since ϕ is non-decreasing, for any $c \in \mathbb{R}$, there exists $c' \in \mathbb{R}$ such that $\phi^{-1}((-\infty, c]) = (-\infty, c']$. It follows from the definition of sigma-algebras generated by random variables that $\sigma(\phi(X_0), \ldots, \phi(X_n)) \subset \sigma(X_0, \ldots, X_n)$.□

It follows from the claim and Proposition 13.2.7 that

$$\mathbf{E}(\phi(X_{n+1})|\sigma(\phi(X_0), \phi(X_1), \ldots, \phi(X_n))) = \mathbf{E}(\mathbf{E}(\phi(X_{n+1})|\sigma(X_0, X_1, \ldots, X_n))|\sigma(\phi(X_0), \phi(X_1), \ldots, \phi(X_n))),$$

with probability 1.
Additionally,

$$\begin{aligned} \mathbf{E}(\phi(X_{n+1})|\sigma(X_0, X_1, \ldots, X_n) &\geq \phi((\mathbf{E}(X_{n+1}|\sigma(X_0, X_1, \ldots, X_n)) && \text{(Conditional Jensen's Inequality)} \\ &\geq \phi(X_n) && (X_n \text{ is a submartingale}) \end{aligned}$$

Therefore,

$$\mathbf{E}(\phi(X_{n+1})|\sigma(\phi(X_0), \phi(X_1), \ldots, \phi(X_n))) \geq \mathbf{E}(\phi(X_n)|\sigma(\phi(X_0), \phi(X_1), \ldots, \phi(X_n))) = \phi(X_n),$$

with probability 1.

(b) It is sufficient to consider non-decreasing and convex functions $\phi_1(x) = \max(x, a)$ and $\phi_2(x) = x^2$ in Exercise 14.4.2, and Exercise 14.4.3, respectively. □

Exercise 14.4.6. Let $\{X_n\}$ be a stochastic process, let τ and ρ be two non-negative-integer valued random variables , and let $m \in \mathbb{N}$.
(a) Prove that τ is a stopping time for $\{X_n\}$ if and only if $\{\tau \leq n\} \in \sigma(X_0, ..., X_n)$ for all $n \geq 0$.
(b) Prove that if τ is a stopping time, then so is $\min(\tau, m)$.
(c) Prove that if τ and ρ are stopping times for $\{X_n\}$, then so is $\min(\tau, \rho)$.

Solution. (a) Let τ satisfies $\{\tau \leq n\} \in \sigma(X_0, ..., X_n)$ for all $n \geq 0$. Then,

$$\{\tau = n\} = \{\tau \leq n\} \bigcap \{\tau \leq n-1\}^c \in \sigma(X_0, ..., X_n),$$

for all $n \geq 0$. Thus, τ is a stopping time.
Conversely, let τ be a stopping time and let $n \geq 0$ be given. Then,

$$\{\tau = m\} \in \sigma(X_0, ..., X_n)$$

for all $0 \leq m \leq n$, and , consequently,

$$\{\tau \leq n\} = \bigcup_{m=0}^{n} \{\tau = m\} \in \sigma(X_0, ..., X_n).$$

(b) First, $\min(\tau, m)$ is a non-negative integer-valued random variable. Second, let $n \geq 0$ be given. Then,

$$\{\min(\tau, m) \leq n\} = \{\tau \leq n\} \bigcup \{m \leq n\},$$

and depending on whether $m \leq n$ or $m > n$ we have that $\{m \leq n\} = \Omega$ or $\{m \leq n\} = \phi$, respectively. Therefore, $\{\min(\tau, m) \leq n\} \in \sigma(X_0, ..., X_n)$. Now, the assertion follows by part (a).

(c) First, since τ and ρ are non-negative-integer valued random variables, $\min(\tau, \rho)$ has the same properties. Second, let $n \geq 0$ be given. Then, by part (a), $\{\tau \leq n\}, \{\rho \leq n\} \in \sigma(X_0, ..., X_n)$ and, consequently,

$$\{\min(\tau, \rho) \leq n\} = \{\tau \leq n\} \bigcup \{\rho \leq n\} \in \sigma(X_0, ..., X_n).$$

Finally, another application of part (a) proves the desired result.□

Exercise 14.4.8. Let $\{X_n\}$ be simple symmetric random walk, with $X_0 = 0$. Let $\tau = \inf\{n \geq 5 : X_{n+1} = X_n + 1\}$ be the first time after 4 which is just before the chain increases. Let $\rho = \tau + 1$.
(a) Is τ a stopping time? Is ρ a stopping time?
(b) Use Theorem 14.1.5., to compute $E(X_\rho)$.
(c) Use the result of part (b) to compute $E(X_\tau)$. Why does this not contradict Theorem 14.1.5.?

Solution. (a) No. τ is not a stopping time because

$$\begin{aligned} \{\tau = n\} &= \{X_6 \neq X_5 + 1, ..., X_n \neq X_{n-1} + 1, X_{n+1} = X_n + 1\} \\ &\notin \sigma(X_0, ..., X_n)\} \end{aligned}$$

for all $n \geq 6$.
Yes. First, $\rho = \tau + 1$ is a non- negative integer-valued random variable. Second, $\{\rho = n\} = \{\tau = n-1\} = \phi$ if $0 \leq n \leq 5$, $\{X_6 \neq X_5 + 1, ..., X_{n-1} \neq X_{n-2} + 1, X_n = X_{n-1} + 1\}$ if $n \geq 6$, implying that $\{\tau = n\} \in \sigma(X_0, ..., X_n)$.

(b) $E(X_\rho) = E(X_0) = 0$.

(c) By definition of τ, we have $X_\rho = X_\tau + 1$ and consequently, $E(X_\tau) = E(X_\rho) - 1 = -1$. This result does not contradict Theorem 14.1.5. due to the fact that τ is not a stopping time and, consequently, the assumption of that Theorem is violated. □

Exercise 14.4.10. Let $0 < a < c$ be integers. Let $\{X_n\}$ be simple symmetric random walk, started at $X_0 = a$. Let $\tau = \inf\{n \geq 1 : X_n = 0, orc\}$.
(a) Prove that $\{X_n\}$ is a martingale.
(b) Prove that $E(X_\tau) = a$.
(c) Use this fact to derive an alternative proof of the gambler's ruin formula given in Section 7.2, for the case $p = 1/2$.

Solution. (a) First, let $Y_n = X_n - a (n \in \mathbb{N})$. Then, by equation 14.0.2, it is a martingale. Hence,

$$E(|X_n|) \leq E(|Y_n|) + |a| < \infty,$$

for all $n \geq 0$. Second, considering the definition of simple random walk in page 75, it follows that:

$$\begin{aligned} E(X_{n+1}|\sigma(X_0,...,X_n)) &= E(X_n + Z_{n+1}|\sigma(X_0,...,X_n)) \\ &= X_n + E(Z_{n+1}|\sigma(X_0,...,X_n)) \\ &= X_n + E(Z_{n+1}|\sigma(Z_0,...,Z_n)) \\ &= X_n + E(Z_{n+1}) \\ &= X_n. \end{aligned}$$

for all $n \geq 0$.

(b) By $\tau = \min(\tau_0, \tau_c)$ and $P(\tau_0 < \infty) = 1 = P(\tau_c < \infty)$, it follows that;

$$P(\tau < \infty) = P(\tau_0 < \infty \cup \tau_c < \infty) = 1.$$

Next, $|X_n|1_{n\leq\tau} \leq c1_{n\leq\tau}$ for all $n \geq 0$. Consequently, by Corollary 14.1.7, it follows that $E(X_\tau) = E(X_0) = a$.

(c) By part (b),

$$a = E(X_\tau) = 0.P(X_\tau = 0) + c.P(X_\tau = c) = cP(\tau_c < \tau_0) = cS(a),$$

implying $S(a) = \frac{a}{c}$. $\square$

Exercise 14.4.12. Let $\{S_n\}$ and τ be as in Example 14.1.13.
(a) Prove that $E(\tau) < \infty$.
(b) Prove that $S_\tau = -\tau + 10$.

Solution. First, we claim that

$$P(\tau > 3m) \leq (\frac{7}{8})^m,$$

for all $m \in \mathbb{N}$. To prove it we use induction. Let $m = 1$, then

$$P(\tau > 3) = 1 - P(\tau = 3) = 1 - \frac{1}{8} = \frac{7}{8}.$$

Assume the assertion holds for positive integer $m > 1$. Then,

$$\begin{aligned} P(\tau > 3(m+1)) &= P(\tau > 3m \cap (r_{3m+1}, r_{3m+2}, r_{3m+3}) \neq (1,0,1)) \\ &= P(\tau > 3m)P((r_{3m+1}, r_{3m+2}, r_{3m+3}) \neq (1,0,1)) \\ &\leq (\frac{7}{8})^m.(\frac{7}{8}) \\ &= (\frac{7}{8})^{m+1}, \end{aligned}$$

proving the assertion for positive integer $m+1$. Second, using Proposition 4.2.9, we have that

$$\begin{aligned} E(\tau) &= \sum_{k=0}^{\infty} P(\tau > k) \\ &= \sum_{m=1}^{\infty} (P(\tau > 3m-1) + P(\tau > 3m-2) + P(\tau > 3m-3)) \\ &\leq 3\sum_{m=1}^{\infty} P(\tau > 3m-3) \\ &= 3 + 3\sum_{m=1}^{\infty} P(\tau > 3m) \\ &\leq 3 + 3\sum_{m=1}^{\infty} (\frac{7}{8})^m \\ &< \infty. \end{aligned}$$

(b) We consider $\tau - 2$ different players . Each of those numbered 1 to $\tau - 3$ has bet and lost \$1 (The person has lost the \$1 bet or, has won the \$1 bet and then has lost the \$2 bet or, has won the \$1 bet and the next \$2 bet but has lost the \$4 bet. In each of these three different cases, the person has totally lost \$1.). The person numbered $\tau - 2$ has won all of \$1, \$2, and \$4 bets successively and has totally won \$7. Accordingly,

$$S_\tau = (\tau - 3)(-1) + 1.7 = -\tau + 10.$$

□

Exercise 14.4.14. Why does the proof of Theorem 14.1.1 fail if $M = \infty$?

Solution. Let $\{X_n\}_{n=0}^{\infty}$ be defined by $X_0 = 0$ and $X_n = \sum_{i=1}^{n} Z_i$ where $\{Z_i\}$ are i.i.d, with $P(Z_i = +1) = P(Z_i = -1) = \frac{1}{2}$ for all $i \in \mathbb{N}$. Then, by Exercise 13.4.2, it is a martingale. Now, by Exercise 4.5.14(c),

$$E(\sum_{i=1}^{\infty} Z_i) \neq \sum_{i=1}^{\infty} E(Z_i),$$

showing that the Proof of Theorem 14.1.1 fails for the case $M = \infty$.□

Exercise 14.4.16. Let $\{X_n\}$ be simple symmetric random walk, with $X_0 = 10$. Let $\tau = \min\{n \geq 1 : X_n = 0\}$, and let $Y_n = X_{\min(n,\tau)}$. Determine (with explanation) whether each of the following statements is true or false.
(a) $E(X_{200}) = 10$.
(b) $E(Y_{200}) = 10$.
(c) $E(X_\tau) = 10$.
(d) $E(Y_\tau) = 10$.
(e) There is a random variable X such that $\{X_n\} \to X$ a.s.
(f) There is a random variable Y such that $\{Y_n\} \to Y$ a.s.

Solution. (a) True. Since $\{X_n\}$ is a martingale, using equation 14.0.4. it follows that $E(X_{200}) = E(X_0) = 10$.

(b) True. Consider the stopping time $\rho = \min(\tau, 200)$. Then, $0 \leq \rho \leq 200$ and, therefore, by Corollary 14.1.3,

$$E(Y_{200}) = E(X_\rho) = E(X_0) = 10.$$

(c) False. Since $P(X_\tau = 0) = 1$, it follows that $E(X_\tau) = 0 \neq 10$.

(d) False. Since $Y_{\tau(w)}(w) = X_{\min(\tau(w),\tau(w))}(w) = X_{\tau(w)}(w)$ for all $w \in \Omega$, it follows that $E(Y_\tau) = E(X_\tau) = 0 \neq 10$.

(e) False. Since $\limsup_n X_n = +\infty$ and $\liminf_n X_n = -\infty$, there is no finite random variable X with $\lim_n X_n = X$ a.s.

(f) True. For the non-negative martingale $\{Y_n\}$ an application of Corollary 14.2.2, yields existence of a finite random variable Y with $\lim_n Y_n = Y$ a.s.□

Exercise 14.4.18. Let $0 < p < 1$ with $p \neq 1/2$, and let $0 < a < c$ be integers. Let $\{X_n\}$ be simple random walk with parameter p , started at $X_0 = a$. Let $\tau = \inf\{n \geq 1 : X_n = 0, orc\}$.
(a) Compute $E(X_\tau)$ by direct computation.
(b) Use Wald's theorem part (a) to compute $E(\tau)$ in terms of $E(X_\tau)$.
(c) Prove that the game's expected duration satisfies

$$E(\tau) = (a - c[((1-p)/p)^a - 1]/[((1-p)/p)^c - 1])/(1-2p).$$

(d) Show that the limit of $E(\tau)$ as $p \to 1/2$ is equal to $a(c-a)$.

Solution. (a)

$$E(X_\tau) = c.P(X_\tau = c) + 0.P(X_\tau = 0) = c.(\frac{1-(\frac{1-p}{p})^a}{1-(\frac{1-p}{p})^c}).$$

(b) From $E(X_\tau) = a + (p-q)E(\tau)$ it follows that

$$E(\tau) = \frac{1}{q-p}(a - E(X_\tau)).$$

(c) By parts (a) and (b), it follows that

$$E(X_\tau) = \frac{1}{1-2p}(a - c.(\frac{1-(\frac{1-p}{p})^a}{1-(\frac{1-p}{p})^c})).$$

(d) Applying I.Hopital's rule, it follows that :

$$\begin{aligned}
\lim_{p\to 1/2} E(\tau) &= \lim_{p\to 1/2} \frac{a(1-(\frac{1-p}{p})^c) - c(1-(\frac{1-p}{p})^a)}{(1-2p)(1-(\frac{1-p}{p})^c)} \\
&=^{Hop} \lim_{p\to 1/2} \frac{-a.c(\frac{1-p}{p})^{c-1}(\frac{-1}{p^2}) - c(-a)(\frac{1-p}{p})^{a-1}(\frac{-1}{p^2})}{-2(1-(\frac{1-p}{p})^c) + (1-2p)(-c(\frac{1-p}{p})^{c-1}(\frac{-1}{p^2}))} \\
&= \lim_{p\to 1/2} ac.\frac{(\frac{1-p}{p})^{a-1} - (\frac{1-p}{p})^{c-1}}{2p^2(1-(\frac{1-p}{p})^c) - (1-2p)c(\frac{1-p}{p})^{c-1}} \\
&=^{Hop} \lim_{p\to 1/2} ac.\frac{(a-1)(\frac{1-p}{p})^{a-2}(\frac{-1}{p^2}) - (c-1)(\frac{1-p}{p})^{c-2}(\frac{-1}{p^2})}{4p(1-(\frac{1-p}{p})^c) + 2p^2(-c(\frac{1-p}{p})^{c-1}(\frac{-1}{p^2})) + 2c(\frac{1-p}{p})^{c-1} - (1-2p)c(c-1)(\frac{1-p}{p})^{c-2}(\frac{-1}{p^2})} \\
&= ac.\frac{(a-1)(-4) - (c-1)(-4)}{0 + 2c + 2c - 0} \\
&= a(c-a).
\end{aligned}$$

□

Exercise 14.4.20. Let $\{X_n\}$ be a martingale with $|X_{n+1} - X_n| \leq 10$ for all n. Let $\tau = \inf\{n \geq 1 : |X_n| \geq 100\}$.
(a) Prove or disprove that this implies that $P(\tau < \infty) = 1$.
(b) Prove or disprove that this implies there is a random variable X with $\{X_n\} \to X$ a.s.
(c) Prove or disprove that this implies that

$$P(\tau < \infty, \text{or there is a random variable X with} \{X_n\} \to X) = 1.$$

Solution. (a) No. As a counterexample, let $X_0 = 0$ and $X_n = \sum_{i=1}^{n} Z_i$ where $\{Z_i\}$ are i.i.d with $P(Z_i = 2^{-i}) = \frac{1}{2}$ and $P(Z_i = -2^{-i}) = \frac{1}{2}$ for all $i \in \mathbb{N}$. Since, $E(Z_i) = 0$ for all $i \in \mathbb{N}$, by Exercise 13.4.12, $\{X_n\}$ is a martingale. Clearly, $|X_{n+1} - X_n| < 2 < 10$ for all $n \in \mathbb{N}$. But,

$$|X_n| = |\sum_{i=1}^{n} Z_i| \leq \sum_{i=1}^{n} |Z_i| < 2$$

for all $n \in \mathbb{N}$, implying $\tau = \inf \phi = \infty$, and consequently, $P(\tau < \infty) = 0 \neq 1$.

(b) No. As a counterexample, consider the simple symmetric random walk with $X_0 = 0$. In this case, $|X_{n+1} - X_n| < 2 \leq 10$ for all n, however, $\limsup_n X_n = +\infty$ and $\liminf_n X_n = -\infty$, showing that there is no finite random variable X with $\lim_n X_n = X$ a.s.
(c) Yes. Define :

$$A^* = \{\limsup_n X_n = -\liminf_n X_n = +\infty\},$$

and

$$B = \{\text{there is a random variable X with} \{X_n\} \to X\}.$$

We claim that $P(A^* \cup B) = 1$. To do this, let $a \in \mathbb{N}$ and define $\rho = \inf\{n \geq 1 : X_n \geq a\}$.
It is stated that $X_{\min(\rho,n)}$ is a martingale, but that has not been proved yet.
Claim. $\{X_{\min(\rho,n)}\}$ is a martingale.
Note that

$$\begin{aligned}
X_{\min(\tau,n+1)} &= \sum_{i=1}^{n} X_i 1_{\tau=i} + X_{n+1} 1_{\tau \geq n+1} \\
&= \sum_{i=1}^{n} X_i 1_{\tau=i} + X_{n+1} 1_{\tau \geq n+1} + X_n 1_{\tau \geq n+1} - X_n 1_{\tau \geq n+1} \\
&= \sum_{i=1}^{n-2} X_i 1_{\tau=i} + X_n 1_{\tau=n} + X_{n+1} 1_{\tau \geq n+1} + X_n 1_{\tau \geq n+1} - X_n 1_{\tau \geq n+1} \\
&= \sum_{i=1}^{n-2} X_i 1_{\tau=i} + (X_n 1_{\tau=n} + X_n 1_{\tau \geq n+1}) + (X_{n+1} 1_{\tau \geq n+1} - X_n 1_{\tau \geq n+1}) \\
&= \sum_{i=1}^{n-2} X_i 1_{\tau=i} + X_n 1_{\tau \geq n} + (X_{n+1} - X_n)\, 1_{\tau \geq n+1} \\
&= X_{\min(\tau,n)} + (X_{n+1} - X_n)\, 1_{\tau \geq n+1} \\
&= X_{\min(\tau,n)} + (X_{n+1} - X_n)\, (1 - 1_{\tau \leq n})
\end{aligned}$$

It follows that

$$\begin{aligned}\mathbb{E}[X_{\min(\tau,n+1)}|X_1,\ldots,X_n] &= \mathbb{E}[X_{\min(\tau,n)} + (X_{n+1}-X_n)\,(1-1_{\tau\leq n})|X_1,\ldots,X_n] \\ &= X_{\min(\tau,n)} + \mathbb{E}[X_{n+1}-X_n|X_1,\ldots,X_n] - \mathbb{E}[(X_{n+1}-X_n)\,1_{\tau\leq n}|X_1,\ldots,X_n] \\ &= X_{\min(\tau,n)} + X_n - X_n - 1_{\tau\leq n}(X_n - X_n) \\ &= X_{\min(\tau,n)}.\end{aligned}$$

By the claim of Exercise 14.4.4 (a), $\sigma(X_{\min(\tau,1)},\ldots,X_{\min(\tau,n)}) \subset \sigma(X_1,\ldots,X_n)$. Hence, by Remark 14.0.1,

$$\mathbb{E}[X_{\min(\tau,n+1)}|\sigma(X_{\min(\tau,1)},\ldots,X_{\min(\tau,n)})] = X_{\min(\tau,n)}.$$

Then, $\{X_{\min(\rho,n)}\}$ is a martingale satisfying :

$$\sup_n E(X^+_{\min(\rho,n)}) \leq a + E(\sup_n |X_{n+1}-X_n|) \leq a + 10 < \infty.$$

Thus, by Theorem 14.2.1, $\{X_{\min(\rho,n)}\}$ converges a.s. Hence, $\{X_n\}$ converges a.s. on $\{\rho = \infty\} = \{\sup_n X_n < a\}$. Let $a \to \infty$, then it follows that $\{X_n\}$ converges a.s. on $\{\sup_n X_n < \infty\}$. A symmetric argument applied to $\{-X_n\}$ shows that $\{X_n\}$ converges a.s. on $\{\inf_n X_n > -\infty\}$, proving our claim. Now, since $\{\tau < \infty\} \supseteq A^*$, the desired result is proved.$\square$

Chapter 15

General stochastic processes

Exercise 15.1.6. Let $X_1, X_2, ...$ be independent, with $X_n \sim \mu_n$.
(a) Specify the finite dimensional distributions $\mu_{t_1,t_2,...,t_k}$ for distinct non-negative integers $t_1 < t_2 < ... < t_k$.
(b) Prove that these $\mu_{t_1,t_2,...,t_k}$ satisfy (15.1.1) and (15.1.2).
(c) Prove that Theorem 7.1.1 follows from Theorem 15.1.3.

Solution. (a) Since the σ-algebra of k-dimensional Borel sets is generated by the class of all bounded rectangles $H_R = \prod_{n=1}^{k} I_n$ where $I_n = (a_n, b_n], (1 \leq n \leq k)$ it is sufficient to specify the distribution $\mu_{t_1,t_2,...,t_k}$ on H_Rs. Now, we have that:

$$\begin{aligned}
\mu_{t_1,t_2,...,t_k}(\prod_{n=1}^{k} I_n) &= P(X_{t_n} \in I_n : 1 \leq n \leq k) \\
&= \prod_{n=1}^{k} P(X_{t_n} \in I_n) \\
&= \prod_{n=1}^{k} \mu_{t_n}(I_n). (\star)
\end{aligned}$$

(b) To prove (15.1.1), several consecutive applications of Proposition 3.3.1 on the both sides of the equality $(\star)$ in part (a) imply that:

$$\begin{aligned}
\mu_{t_1,t_2,...,t_k}(\prod_{n=1}^{k} H_n) &= \prod_{n=1}^{k} \mu_{t_n}(H_n) \\
&= \prod_{n=1}^{k} \mu_{t_{s(n)}}(H_{s(n)}) \\
&= \mu_{t_{s(1)},t_{s(2)},...,t_{s(k)}}(\prod_{n=1}^{k} H_{s(n)}),
\end{aligned}$$

for all Borel $H_1, H_2, ..., H_k \subseteq \mathbb{R}$ and all permutations $(s(n))_{n=1}^k$ of $(n)_{n=1}^k$.
To prove (15.1.2) , by a similar argument used for the proof of (15.1.1), it follows that :

$$\begin{aligned} \mu_{t_1,t_2,...,t_k}(H_1 \times H_2 \times ... \times H_{k-1} \times \mathbb{R}) &= \prod_{n=1}^{k-1} \mu_{t_n}(H_n)\mu_{t_k}(\mathbb{R}) \\ &= \prod_{n=1}^{k-1} \mu_{t_n}(H_n) \\ &= \mu_{t_1,t_2,...,t_{k-1}}(H_1 \times H_2 \times ... \times H_{k-1}), \end{aligned}$$

for all Borel $H_1, H_2, ..., H_{k-1} \subseteq \mathbb{R}$.

(c) We consider the family of Borel probability measures $\{\mu_{t_1,t_2,...,t_k} : k \in \mathbb{N}, t_i \in \mathbb{N}\text{distinct}\}$ with $\mu_{t_1,t_2,...,t_k}$ defined by

$$\mu_{t_1,t_2,...,t_k}(\prod_{n=1}^{k} H_n) = \prod_{n=1}^{k} \mu_{t_n}(H_n),$$

for all Borel $H_1, H_2, ..., H_k \subseteq \mathbb{R}$. Then, by part (b) it satisfies the consistency conditions (C1) and (C2). Consequently, by Theorem 15.1.3, there is a probability space $(\mathbb{R}^{\mathbb{N}}, \mathcal{F}^{\mathbb{N}}, P)$ and random variables $\{X_n\}$ defined on that triple such that

$$\mu_{t_1,t_2,...,t_k}(H) = P((X_{t_1}, X_{t_2}, ..., X_{t_k}) \in H), \qquad (\star\star)$$

for all $k \in \mathbb{N}$, distinct $t_1, ..., t_k \in \mathbb{N}$ and Borel $H \subseteq \mathbb{R}^k$. Now, put

$$H = \mathbb{R} \times ... \times \mathbb{R} \times H_n \times \mathbb{R} ... \times \mathbb{R},$$

where $H_n \subseteq \mathbb{R}$ is Borel $(1 \leq n \leq k)$. Then, by $(\star\star)$:

$$\mu_{t_n}(H_n) = P(X_{t_n} \in H_n) = \mathcal{L}_{t_n}(H_n),$$

for all $n \in \mathbb{N}$. $\square$

Exercise 15.2.4. Consider a Markov chain which is ϕ irreducible with respect to some non-zero σ-finite measure ψ, and which is periodic with corresponding disjoint subsets $\mathcal{X}_1, ..., \mathcal{X}_d$. Let $B = \cup_i \mathcal{X}_i$.
(a) Prove that $P^n(x, B^c) = 0$ for all $x \in B$.
(b) Prove that $\psi(B^c) = 0$.
(c) Prove that $\psi(\mathcal{X}_i) > 0$ for some i.

Solution. (a) We prove the assertion by induction. Let $n = 1$ and $x \in B = \cup_{i=1}^d \mathcal{X}_i$. Then, there exists an unique $1 \leq i \leq d$ such that $x \in \mathcal{X}_i$, and consequently, $P(x, \mathcal{X}_{i+1}) = 1$ if $1 \leq i \leq d-1$ or $P(x, \mathcal{X}_1) = 1$ if $i = d$. Hence,

$$0 \leq P(x, B^c) \leq P(x, \mathcal{X}_{i+1}^c) = 0 \text{ if } 1 \leq i \leq d-1$$

or

$$0 \leq P(x, B^c) \leq P(x, \mathcal{X}_1^c) = 0 \text{ if } i = d,$$

implying $P^1(x, B^c) = 0$. Next, assume the assertion holds for positive integer $n > 1$. Then, by $P^n(z, B^c) = 0 (z \in B^c)$ it follows that:

$$\begin{aligned} P^{n+1}(x, B^c) &= \int_{\mathcal{X}} P(x, dz) P^n(z, B^c) \\ &= \int_B P(x, dz) P^n(z, B^c) + \int_{B^c} P(x, dz) P^n(z, B^c) \\ &= 0, \end{aligned}$$

proving the assertion for positive integer $n + 1$.

(b) Since $\tau_{B^c} = \inf\{n \geq 0 : X_n \in B^c\} = \inf \phi = \infty$, it follows that $P_x(\tau_{B^c} < \infty) = 0$, for all $x \notin B^c$, and, consequently, $\psi(B^c) = 0$.

(c) Using part (b), we have that:

$$\sum_{i=1}^{d} \psi(\mathcal{X}_i) = \psi(B) = \psi(\mathcal{X}) - \psi(B^c) = \psi(\mathcal{X}) > 0,$$

implying that $\psi(\mathcal{X}_i) > 0$ for some $1 \leq i \leq d$. $\square$

Exercise 15.2.6. (a) Prove that a Markov chain on a countable state space $\mathcal{X}$ is ϕ-irreducible if and only if there is $j \in \mathcal{X}$ such that $P_i(\tau_j < \infty) > 0$, for all $i \in \mathcal{X}$.
(b) Give an example of a Markov chain on a countable state space which is ϕ- irreducible, but which is not irreducible in the sense of Subsection 8.2.

Solution. (a) Let the given Markov chain on the countable state space $\mathcal{X}$ be ϕ irreducible. Then, for some $A = \{j\} \in \mathcal{F}$ with $\psi(A) > 0$, we have $\tau_A = \tau_j$ implying $P_i(\tau_j < \infty) > 0$ for all $i \in \mathcal{X}$. Conversely, let there is $j \in \mathcal{X}$ such that $P_i(\tau_j < \infty) > 0$ for all $i \in \mathcal{X}$. Then, we consider the σ-finite measure ψ on $\mathcal{F}$ defined by:

$$\psi(A) = \delta_j(A),$$

for all $A \in \mathcal{F}$. Now, let $A \in \mathcal{F}$ with $\psi(A) > 0$, then $j \in A$ implying that $\tau_A \leq \tau_j$ and consequently, $P_i(\tau_A < \infty) \geq P_i(\tau_j < \infty) > 0$ for all $i \in \mathcal{X}$.

(b) Consider the Markov chain given in the solution of Exercise 8.5.20(a), which is reducible in the sense of Subsection 8.2. However, for $j = 2$ we have that:

$$\begin{aligned} P_i(\tau_2 < \infty) &= P(\exists n \geq 0 : X_n = 2 | X_0 = i) \\ &\geq \sum_{i_1, \ldots, i_{n-1}} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-2} i_{n-1}} p_{i_{n-2} 2} \\ &> 0, \end{aligned}$$

for $i = 1, 2$. Hence, by part (a), this Markov chain is ϕ-irreducible.$\square$

Exercise 15.2.8. Consider a discrete Markov chain with state space $\mathcal{X} = \mathbb{R}$, and with transition probabilities such that $P(x, \cdot)$ is uniform on the interval $[x - 1, x + 1]$. Determine whether or not this chain is ϕ-irreducible.

Solution. Yes, this chain is ϕ-irreducible. Indeed, $P(x, \cdot)$ has positive density (with respect to Lebesgue measure) throughout $[x - 1, x + 1]$. Then, by the convolution formula, $P^2(x, \cdot)$ has positive density throughout $[x - 2, x + 2]$. By induction, $P^n(x, \cdot)$ has positive density throughout $[x - n, x + n]$. Now, if $A \subseteq \mathcal{X}$ has positive Lebesgue measure, then $A \cap [x - n, x + n]$ must have positive Lebesgue measure for some $n \in \mathbf{N}$. It then follows that $P^n(x, A) > 0$. Hence, the chain is ϕ-irreducible where ϕ is Lebesgue measure. $\square$

Exercise 15.2.10. Consider the Markov chain with $\mathcal{X} = \mathbb{R}$, and with $P(x, .) = N(x, 1)$ for each $x \in \mathcal{X}$.
(a) Prove that this chain is ϕ-irreducible and aperiodic.
(b) Prove that his chain does not have a stationary distribution. Relate this to Theorem 15.2.3.

Solution. (a) Here $P(x, \cdot)$ has an everywhere-positive density (with respect to Lebesgue measure). It follows that if $A \subseteq \mathcal{X}$ has positive Lebesgue measure, then $P(x, A) > 0$ for all $x \in \mathbf{X}$. So, the chain is ϕ-irreducible where ϕ is Lebesgue measure. As for aperiodicity, if $\mathcal{X}_1, \ldots, \mathcal{X}_d$ is a periodic decomposition for some $d \geq 2$, then if $x_1 \in \mathcal{X}_1$ then $P(x_1, \mathcal{X}_2) > 0$. This implies that $\mathcal{X}_2$ has positive Lebesgue measure, which in turn implies that $P(x, \mathcal{X}_2) > 0$ for all $x \in \mathcal{X}$, even for $x \in \mathcal{X}_2$. This contradicts the assumption of periodic decomposition. So, no such periodic decomposition exists, i.e. the chain is aperiodic.
(b) This chain is equivalent to adding an independent $N(0, 1)$ random variable at each iteration. So, if a stationary probability distribution $\pi(\cdot)$ existed, it would have to satisfy the property that if $X \sim \pi(\cdot)$ and $Y \sim N(0, 1)$ are independent, then $X + Y \sim \pi(\cdot)$. It would follow by induction that if $Y_1, Y_2, \ldots$ are i.i.d. $\sim N(0, 1)$ (and independent of X), then $Z_n \equiv X + Y_1 + \ldots + Y_n \sim \pi(\cdot)$ for all n, which would imply that for any $a \in \mathbf{R}$, we have $\pi((a, \infty)) = \lim_{n\to\infty} P(Z_n > a) = 1/2$. This is impossible, since we must have $\lim_{a\to\infty} \pi((a, \infty)) = 0$. Hence, no such stationary probability distribution exists. Hence, Theorem 15.2.3 does not apply, and the distributions $P^n(x, \cdot)$ need not converge. (In fact, $\lim_{n\to\infty} P^n(x, A) = 0$ for every bounded subset A.) $\square$

Exercise 15.2.12. Show that the finite-dimensional distributions implied by (15.2.1) satisfy the two consistency conditions of the Kolmogorov Existence Theorem. What does this allow us to conclude?

Solution. Since (15.2.1) specifies the probabilities for random variables specifically in the order $X_0, X_1, \ldots, X_n$, the probabilities for random variables in any other order would be found simply by un-permuting them and then applying the same formula (15.2.1). Hence, (C1) is immediate. Similarly, since (15.2.1) specifies the probabilities for the entire sequence $X_0, X_1, \ldots, X_n$, the probabilities for just a subset of these variables would be found by integrating over the missing variables, i.e. by setting $A_i = \mathbf{R}$ for each of them, thus automatically satisfying (C2). Hence, since (C1) and (C2) are satisfied, it follows that there must exist some probability triple on which random variables can be defined which satisfy the probability specifications (15.2.1). Informally speaking, this says that Markov chains on general state spaces defined by (15.2.1) must actually exist. □

Exercise 15.3.6. Let $\mathcal{X} = \{1, 2\}$, and let $Q = (q_{ij})$ be the generator of a continuous time Markov process on $\mathcal{X}$, with

$$Q = \begin{pmatrix} -3 & 3 \\ 6 & -6 \end{pmatrix}.$$

Compute the corresponding transition probabilities $P^t = (p^t_{ij})$ of the process, for any $t > 0$.

Solution. For given generator matrix Q, an easy argument by induction shows that (Check!):

$$Q^m = \begin{pmatrix} 3^{2m-1}(-1)^m & -3^{2m-1}(-1)^m \\ -2.3^{2m-1}(-1)^m & 2.3^{2m-1}(-1)^m \end{pmatrix} \qquad m \geq 1.$$

Now, by Exercise 15.3.5, it follows that:

$$\begin{aligned} P^t &= \exp(tQ) \\ &= I + \sum_{m=1}^{\infty} \frac{t^m}{m!} Q^m \\ &= I + \sum_{m=1}^{\infty} \frac{t^m}{m!} \begin{pmatrix} 3^{2m-1}(-1)^m & -3^{2m-1}(-1)^m \\ -2.3^{2m-1}(-1)^m & 2.3^{2m-1}(-1)^m \end{pmatrix} \\ &= \begin{pmatrix} 1 + \frac{1}{3}\sum_{m=1}^{\infty} \frac{(-9t)^m}{m!} & -\frac{1}{3}\sum_{m=1}^{\infty} \frac{(-9t)^m}{m!} \\ \frac{-2}{3}\sum_{m=1}^{\infty} \frac{(-9t)^m}{m!} & 1 + \frac{2}{3}\sum_{m=1}^{\infty} \frac{(-9t)^m}{m!} \end{pmatrix} \\ &= \begin{pmatrix} 1 + \frac{1}{3}(e^{-9t} - 1) & \frac{-1}{3}(e^{-9t} - 1) \\ \frac{-2}{3}(e^{-9t} - 1) & 1 + \frac{2}{3}(e^{-9t} - 1). \end{pmatrix} \qquad (t > 0) \end{aligned}$$

□

Exercise 15.3.8. For a Markov chain on a finite state space $\mathcal{X}$ with generator Q, prove that $\{\pi_i\}_{i\in\mathcal{X}}$ is a stationary distribution if and only if $\pi Q = 0$, i.e. if and only if $\sum_{i\in\mathcal{X}} \pi_i q_{ij} = 0$ for all $j \in \mathcal{X}$.

Solution. First, let $\{\pi_i\}_{i\in\mathcal{X}}$ be a stationary distribution. Then, by equation 15.3.4 and finiteness of $\mathcal{X}$ it follows that:

$$\begin{aligned} \sum_{i\in\mathcal{X}} \pi_i q_{ij} &= \sum_{i\in\mathcal{X}} \pi_i (\lim_{t\to 0^+} \frac{p^t_{ij} - \delta_{ij}}{t}) \\ &= \lim_{t\to 0^+} \frac{\sum_{i\in\mathcal{X}} \pi_i p^t_{ij} - \sum_{i\in\mathcal{X}} \pi_i \delta_{ij}}{t} \\ &= \lim_{t\to 0^+} \frac{\pi_j - \pi_j}{t} \\ &= 0, \end{aligned}$$

for all $j \in \mathcal{X}$, i.e. $\pi Q = 0$.
Conversely, assume $\pi Q = 0$, then by Exercise 15.3.5 we have that:

$$\begin{aligned} \pi P^t &= \pi \exp(t.Q) \\ &= \pi(I + \sum_{m=1}^{\infty} \frac{t^m}{m!} Q^m) \\ &= \pi + \sum_{m=1}^{\infty} \frac{t^m}{m!} (\pi Q) Q^{m-1} \\ &= \pi, \end{aligned}$$

for all $t > 0$. Hence, π is a stationary distribution.□

Exercise 15.3.10. (*Poisson Process.*) Let $\lambda > 0$, let $\{Z_n\}$ be i.i.d. $\sim Exp(\lambda)$, and let $T_n = Z_1 + Z_2 + ... + Z_n$ for $n \geq 1$. Let $\{N_t\}_{t \geq 0}$ be a continuous time Markov process on the state space $\mathcal{X} = \{0, 1, 2, ...\}$, with $N_0 = 0$, which does not move except at the times T_n. (Equivalently, $N_t = \#\{n \in \mathbb{N} : T_n \leq n\}$; intuitively, N_t counts the number of events by time t.)
(a) Compute the generator Q for this process.
(b) Prove that $P(N_t \leq m) = e^{\lambda t}(\lambda t)^m/m! + P(N_t \leq m - 1)$ for $m = 0, 1, 2, ...$
(c) Conclude that $P(N_t = j) = e^{\lambda t}(\lambda t)^j/j!$, i.e. that $N_t \sim Poisson(\lambda t)$.

Solution. (a) Since the exponential distribution has memoryless property, it follows that:

$$\begin{aligned} P(N_{t+\triangle t} \geq i+1 | N_t = i) &= P(T_{i+1} \leq t + \triangle t | T_i \leq t, T_{i+1} > t) \\ &= P(Z_{i+1} \leq \triangle t + t - T_i | T_i \leq t, Z_{i+1} > t - T_i) \\ &= P(Z_{i+1} \leq \triangle t) \\ &= 1 - \exp(-\lambda \triangle t), \end{aligned}$$

for all $i \in \mathcal{X}$. Consequently,

$$q_{i,i+1} = \lim_{\triangle t \to 0^+} \frac{P(N_{t+\triangle t} = i+1 | N_t = i)}{\triangle t} = \lim_{\triangle t \to 0^+} \frac{1 - \exp(-\lambda \triangle t)}{\triangle t} = \lambda,$$

and,

$$q_{i,i} = \lim_{\triangle t \to 0^+} \frac{P(N_{t+\triangle t} = i | N_t = i) - 1}{\triangle t} = -\lambda,$$

for all $i \in \mathcal{X}$. Next, by considering $\sum_{j \in \mathcal{X}} q_{ij} = 0$ for all $i \in \mathcal{X}$, it follows that $\sum_{j \neq i, i+1} q_{ij} = 0$ for all $i \in \mathcal{X}$. On the other hand, $q_{ij} \geq 0$ for all $j \neq i, i+1$ and , consequently, $q_{ij} = 0$ for all $j \neq i, i+1$ and $i \in \mathcal{X}$. To sum up , the generator matrix has the form:

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & ... \\ 0 & -\lambda & \lambda & 0 & ... \\ 0 & 0 & -\lambda & \lambda & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

(b) For the Poisson process $N(t)$ with rate parameter λ, and the mth arrival time $T_m (m \geq 1)$, the number of arrivals before some fixed time t is less than $m + 1$ if and only if the waiting time until the

$m+1$th arrival is more than t. Hence the event $\{N(t) < m+1\} = \{N(t) \leq m\}$ occurs if and only if the event $\{T_{m+1} > t\}$ occurs. Thus, the probabilities of these events are the same:

$$P(N_t \leq m) = P(T_{m+1} > t). \qquad (\star)$$

Next, since $Z_k \sim Exp(\lambda) = \text{Gamma}(1, \lambda), (1 \leq k \leq m)$, by Exercise 9.5.17 and induction it follows that $T_m \sim \text{Gamma}\ (m, \lambda), (m \geq 1)$. Therefore, by equality $(\star)$ and integration by parts it follows that:

$$\begin{aligned}
P(N_t \leq m) &= P(T_{m+1} > t) \\
&= \int_t^\infty \frac{(\lambda x)^m \lambda e^{-\lambda x}}{m!} dx \\
&= \int_t^\infty (\lambda x)^m d(\frac{-e^{-\lambda x}}{m!}) dx \\
&= (\lambda x)^m (\frac{-e^{-\lambda x}}{m!})|_t^\infty + \int_t^\infty \frac{e^{-\lambda x}}{m!} m(\lambda x)^{m-1} \lambda dx \\
&= e^{-\lambda t} \frac{(\lambda t)^t}{m!} + P(T_m > t) \\
&= e^{-\lambda t} \frac{(\lambda t)^t}{m!} + P(N_t \leq m-1),
\end{aligned}$$

for all $m \geq 1$.

(c) By part (b), we have that :

$$P(N_t = j) = P(N_t \leq j) - P(N_t \leq j-1) = e^{\lambda t} (\lambda t)^j / j!,$$

for all $t > 0$ and $j = 0, 1, 2, ...$ □

Exercise 15.3.12. (a) Let $\{N_t\}_{t \geq 0}$ be a Poisson process with rate $\lambda > 0$, let $0 < s < t$, and let U_1, U_2 be i.i.d. $\sim Uniform[0, t]$.
(a) Compute $P(N_s = 0 | N_t = 2)$.
(b) Compute $P(U_1 > s, U_2 > s)$.
(c) Compare the answers to parts (a) and (b). What does this comparison seem to imply?

Solution. (a)

$$\begin{aligned}
P(N_s = 0 | N_t = 2) &= \frac{P(N_t = 2 | N_s = 0) P(N_s = 0)}{P(N_t = 2)} \\
&= \frac{P(N_{t-s} = 2) P(N_s = 0)}{P(N_t = 2)} \\
&= \frac{(\frac{e^{-\lambda(t-s)}(\lambda(t-s))^2}{2!}) e^{-\lambda s}}{\frac{e^{-\lambda t}(\lambda t)^2}{2!}} \\
&= (\frac{t-s}{t})^2.
\end{aligned}$$

(b)

$$\begin{aligned} P(U_1 > s, U_2 > s) &= P(U_1 > s)P(U_2 > s) \\ &= (\int_s^t \frac{dx}{t})(\int_s^t \frac{dx}{t}) \\ &= (\frac{t-s}{t})^2. \end{aligned}$$

(c) By comparing part (a) and part (b), we have that:

$$P(N_s = 0|N_t = 2) = P(U_1 > s, U_2 > s) \qquad 0 < s < t.$$

□

Exercise 15.4.4. Let $\{B_t\}_{t\geq 0}$ be Brownian motion, and let $X_t = 2t + 3B_t$ for $t \geq 0$.
(a) Compute the distribution of X_t for $t \geq 0$.
(b) Compute $E(X_t^2)$ for $t \geq 0$.
(c) Compute $E(X_s X_t)$ for $0 < s < t$.

Solution. (a) If $Y \sim N(\mu, \sigma^2)$, then $aY + b \sim N(a\mu + b, a^2\sigma^2)$, where $a \neq 0, b \in \mathbb{R}$. Hence, from $B_t \sim N(0,t)$ it follows that $X_t \sim N(2t, 9t), t \geq 0$

(b)

$$E(X_t^2) = var(X_t) + E^2(X_t) = 9t + (2t)^2 = 4t^2 + 9t.$$

(c)

$$\begin{aligned} E(X_s X_t) &= E((3B_s + 2s)(3B_t + 2t)) \\ &= E(9B_sB_t + 6tB_s + 6sB_t + 4st) \\ &= 9E(B_sB_t) + 6tE(B_s) + 6sE(B_t) + 4st \\ &= 9s + 4st.(0 \leq s \leq t) \end{aligned}$$

□

Exercise 15.4.6. (a) Let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz function, i.e. a function for which there exists $\alpha \in \mathbb{R}$ such that $|f(x) - f(y)| \leq \alpha|x - y|$ for all $x, y \in \mathbb{R}$. Compute $\lim_{h\to 0^+}(f(t+h) - f(t))^2/h$ for any $t \in \mathbb{R}$.
(b) Let $\{B_t\}$ be Brownian motion. Compute $\lim_{h\to 0^+} E((B_(t+h) - B_(t))^2/h)$ for any $t > 0$.
(c) What do parts (a) and (b) seem to imply about Brownian motion?

Solution. (a) By considering :

$$0 \leq |(f(t+h) - f(t))^2/h| \leq \frac{\alpha^2|t+h-t|^2}{h} = \alpha^2.h,$$

and $\lim_{h\to 0^+} \alpha^2.h = 0$, it follows that, $\lim_{h\to 0^+}(f(t+h) - f(t))^2/h = 0$, for any $t > 0$.

(b) Since $\mathcal{L}(B_{t+h} - B_t) = N(0, t+h-t) = N(0,h) = \mathcal{L}(B_h)$, we have that :

$$\begin{aligned}\lim_{h\to 0^+} E((B_{t+h} - B_t)^2/h) &= \lim_{h\to 0^+} \frac{Var(B_h) + E^2(B_h)}{h}, \\ &= \lim_{h\to 0^+} \frac{h + 0^2}{h}, \\ &= 1.\end{aligned}$$

(c) Let $B_t, t \geq 0$ be Lipschitz function. Then, by part (a), $\lim_{h\to 0^+}(B_{t+h} - B_t)^2/h = 0$, and consequently, by Part (b),

$$0 = \lim_{h\to 0^+} E((B_{t+h} - B_t)^2/h) = 1,$$

a contradiction. Therefore, $B_t, t \geq 0$ is not a Lipschitz function, yielding that at least one of four Dini-derivatives of $B_t, t \geq 0$, defined by :

$$\begin{aligned}D^+ B_t &= \limsup_{h\to 0^+} \frac{B_{t+h} - B_t}{h}, \\ D^- B_t &= \limsup_{h\to 0^+} \frac{B_t - B_{t-h}}{h}, \\ D_+ B_t &= \liminf_{h\to 0^+} \frac{B_{t+h} - B_t}{h}, \\ D_- B_t &= \liminf_{h\to 0^+} \frac{B_t - B_{t-h}}{h},\end{aligned}$$

is unbounded on $[0, \infty)$. □

Exercise 15.5.2. Let $\{X_t\}_{t\in T}$ and $\{X_t^{'}\}_{t\in T}$ be stochastic processes with the countable time index T. Suppose $\{X_t\}_{t\in T}$ and $\{X_t^{'}\}_{t\in T}$ have identical finite-dimensional distributions. Prove or disprove that $\{X_t\}_{t\in T}$ and $\{X_t^{'}\}_{t\in T}$ must have the same full joint distribution.

Solution. We show that if the distributions of both stochastic processes $\{X_t\}_{t\in T}$ and $\{X_t^{'}\}_{t\in T}$ satisfy the "infinite version" of the equation (15.1.2), then the answer is affirmative. For this case, let $T =$

$\{t_n\}_{n=1}^{\infty}$. Then,two applications of Proposition 3.3.1. yield:

$$\begin{aligned} P(\prod_{m=1}^{\infty} X_{t_m} \in \prod_{m=1}^{\infty} H_m) &= P(\lim_{n\to\infty}(\prod_{m=1}^{\infty} X_{t_m} \in H_1 \times ... \times H_n \times \mathbb{R} \times ...)) \\ &= \lim_{n\to\infty} P(\prod_{m=1}^{\infty} X_{t_m} \in H_1 \times ... \times H_n \times \mathbb{R} \times ...) \\ &= \lim_{n\to\infty} P(\prod_{m=1}^{n} X_{t_m} \in H_1 \times ... \times H_n) \\ &= \lim_{n\to\infty} P^{'}(\prod_{m=1}^{n} X^{'}_{t_m} \in H_1 \times ... \times H_n) \\ &= \lim_{n\to\infty} P^{'}(\prod_{m=1}^{\infty} X^{'}_{t_m} \in H_1 \times ... \times H_n \times \mathbb{R} \times ...) \\ &= P^{'}(\lim_{n\to\infty}(\prod_{m=1}^{\infty} X^{'}_{t_m} \in H_1 \times ... \times H_n \times \mathbb{R} \times ...)) \\ &= P^{'}(\prod_{m=1}^{\infty} X^{'}_{t_m} \in \prod_{m=1}^{\infty} H_m) \end{aligned}$$

for all Borel $H_m \subseteq \mathbb{R}(1 \leq m \leq < \infty$. $\square$

Exercise 15.6.8. Let $\{B_t\}_{t\geq 0}$ be standard Brownian motion, with $B_0 = 0$. Let $X_t = \int_0^t a ds + \int_0^t b dB_s = at + bB_t$ be a diffusion with constant drift $\mu(x) = a > 0$ and constant volatility $\sigma(x) = b > 0$. Let $Z_t = \exp[-2aX_t/b^2]$.
(a) Prove that $\{Z_t\}_{t\geq 0}$ is a martingale, i.e. that $E[Z_t|Z_u(0 \leq u \leq s)] = Z_s$ for $0 < s < t$.
(b) Let $A, B > 0$ and let $T_A = \inf\{t \geq 0 : X_t = A\}$ and $T_{-B} = \inf\{t \geq 0 : X_t = -B\}$ denote the first hitting times of A and $-B$, respectively. Compute $P(T_A < T_{-B})$.

Solution. (a) Let $c = -2a/b^2$, so $Z_t = \exp[cX_t]$. Then by the independent increments property,

$$\begin{aligned} E[Z_t|Z_u(0 \leq u \leq s)] &= E[Z_s(Z_t/Z_s)|Z_u(0 \leq u \leq s)] \\ &= E[Z_s \exp(c(X_t - X_s))|Z_u(0 \leq u \leq s)] \\ &= Z_s\, E[\exp(c(X_t - X_s))]\,. \qquad (**) \end{aligned}$$

But $X_t - X_s = a(t-s) + b(B_t - B_s) \sim N(a(t-s),\, b^2(t-s))$, so $X_t - X_s = a(t-s) + b\sqrt{t-s}\,U$ where $U \sim N(0,1)$. Hence,

$$\begin{aligned} E[\exp(c(X_t - X_s))] &= E[\exp(ac(t-s) + bc\sqrt{t-s}\,U)] \\ &= \exp[ac(t-s) + b^2c^2(t-s)/2] \\ &= \exp[a(-2a/b^2)(t-s) + b^2(4a^2/b^4)(t-s)/2] \\ &= \exp[-(2a^2/b^2)(t-s) + (2a^2/b^2)(t-s)] \\ &= \exp[0] \\ &= 1\,. \end{aligned}$$

Substituting this into $(**)$, we have: $E[Z_t|Z_u(0 \leq u \leq s)] = Z_s\,(1) = Z_s$.

(b) Let $p = P(T_A < T_B)$, and let $\tau = \min(T_A, T_{-B})$. Then τ is a stopping time for the martingale. So, by Corollary 14.1.7, we must have $E(Z_\tau) = E(Z_0) = \exp(0+0) = 1$, so that $p\,e^{cA} + (1-p)e^{-cB} = 1$, whence $p(e^{cA} - e^{-cB}) = 1 - c^{-cB}$, so $p = (1 - e^{-cB})/(e^{cA} - e^{-cB})$. $\square$

Exercise 15.6.10. Let $\{p_{ij}^t\}$ be the transition probabilities for a Markov chain on a finite state space $\mathcal{X}$. Define the matrix $Q = (q_{ij})$ by (15.3.4.). Let $f : \mathcal{X} \to \mathbb{R}$ be any function, and let $i \in \mathcal{X}$. Prove that $(Qf)_i$ (i.e., $\sum_{k\in\mathcal{X}} q_{ik} f_k$) corresponds to $(Qf)(i)$ from (15.6.4).

Solution. Using (15.6.7) and finiteness of $\mathcal{X}$, we have that:

$$
\begin{aligned}
(Qf)(i) &= \lim_{t\to 0^+} \frac{E_i(f(X_{t_0+t}) - f(X_{t_0}))}{t} \\
&= \lim_{t\to 0^+} \frac{E(f(X_{t_0+t}) - f(X_{t_0})|X_{t_0} = i)}{t} \\
&= \lim_{t\to 0^+} \frac{\sum_{k\in\mathcal{X}}(p_{ik}^t f_k - \delta_{ik} f_k)}{t} \\
&= \lim_{t\to 0^+} \sum_{k\in\mathcal{X}} (\frac{1}{t}(p_{ik}^t - \delta_{ik}) f_k) \\
&= \sum_{k\in\mathcal{X}} \lim_{t\to 0^+} \frac{1}{t}(p_{ik}^t - \delta_{ik}) f_k \\
&= \sum_{k\in\mathcal{X}} q_{ik} f_k \\
&= (Qf)_i.
\end{aligned}
$$

The finiteness of $\mathcal{X}$ allows us to interchange limit $\lim_{t\to 0^+}$ with sum $\sum_{k\in\mathcal{X}}$ in the 5th equation above. $\square$

Exercise 15.6.12. Suppose a diffusion $\{X_t\}_{t\geq 0}$ has generator given by $(Qf)(x) = \frac{-x}{2} f'(x) + \frac{1}{2} f''(x)$. (Such a diffusion is called an Ornstein-Uhlenbeck process.)
(a) Write down a formula for dX_t.
(b) Show that $\{X_t\}_{t\geq 0}$ is reversible with respect to the standard normal distribution, $N(0, 1)$.

Solution. (a) By equation (15.6.6) it follows that:

$$\mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x) = \frac{-x}{2}f'(x) + \frac{1}{2}f''(x),$$

and, consequently, $\mu(x) = \frac{-x}{2}$ and $\sigma(x) = 1$. Now, by recent result and equation (15.6.2) we conclude that :

$$dX_t = dB_t - \frac{X_t}{2}dt.$$

(b) The standard normal distribution has probability distribution

$$g(x) = \frac{1}{\sqrt{2\pi}}\exp(\frac{-x^2}{2}) \qquad -\infty < x < \infty,$$

and, therefore, $\mu(x) = -\frac{x}{2} = \frac{1}{2}\frac{g'(x)}{g(x)}$. Besides, $\sigma(x) = 1$, and the desired result follows by Exercise 15.6.9.□

Exercise 15.7.2. Consider the Ornstein-Uhlenbeck process $\{X_t\}_{t\geq 0}$ of Exercise 15.6.12, with generator $(Qf)(x) = \frac{-x}{2}f'(x) + \frac{1}{2}f''(x)$.
(a) Let $Y_t = X_t^2$ for each $t \geq 0$. Compute dY_t.
(b) Let $Z_t = X_t^3$ for each $t \geq 0$. Compute dZ_t.

Solution. (a),(b) For $f_n(x) = x^n (n \geq 1)$, $\sigma(x) = 1$, and $\mu(x) = -\frac{x}{2}$ an application of the equation (15.7.1) yields:

$$\begin{aligned} d(f_n(X_t)) &= f_n'(X_t)\sigma(X_t)dB_t + (f_n'(X_t)\mu(X_t) + \frac{1}{2}f_n''(X_t)\sigma^2(X_t))dt \\ &= nX_t^{n-1}dB_t + (\frac{-nX_t^n + n(n-1)X_t^{n-2}}{2})dt. \end{aligned}$$

□

Exercise 15.8.6. Show that (15.8.4) is indeed equal to (15.8.5).

Solution. Using change of variable method and

$$\Phi(-x) = 1 - \Phi(x)$$

for all $x \geq 0$ we have that:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{e^{-rT}\max(0, P_0\exp(\sigma x + (r - \frac{1}{2}\sigma^2)T) - q)e^{-\frac{x^2}{2T}}}{\sqrt{2\pi T}}dx &= \int_{\frac{1}{\sigma}(\log(\frac{q}{P_0}) - (r - \frac{1}{2}\sigma^2)T))}^{\infty} \frac{e^{-rT}(P_0\exp((\sigma x + (r - \frac{1}{2}\sigma^2)T) - q)e^{-\frac{x^2}{2T}}}{\sqrt{2\pi T}}dx \\ &= P_0\int_{\frac{1}{\sigma}(\log(\frac{q}{P_0}) - (r - \frac{1}{2}\sigma^2)T))}^{\infty} \frac{1}{\sqrt{2\pi T}}\exp(\frac{-(x - \sigma T)^2}{2T})dx \\ &- qe^{-rT}\int_{\frac{1}{\sigma}(\log(\frac{q}{P_0}) - (r - \frac{1}{2}\sigma^2)T))}^{\infty} \frac{1}{\sqrt{2\pi T}}\exp(\frac{-x^2}{2T})dx \\ &= P_0\int_{\frac{1}{\sigma\sqrt{T}}(\log(\frac{q}{P_0}) - (r + \frac{1}{2}\sigma^2)T))}^{\infty} \frac{1}{\sqrt{2\pi}}\exp(\frac{-s^2}{2})ds \\ &- qe^{-rT}\int_{\frac{1}{\sigma\sqrt{T}}(\log(\frac{q}{P_0}) - (r - \frac{1}{2}\sigma^2)T))}^{\infty} \frac{1}{\sqrt{2\pi}}\exp(\frac{-s^2}{2})ds \\ &= P_0\Phi(\frac{1}{\sigma\sqrt{T}}(\log(\frac{P_0}{q}) + (r + \frac{1}{2}\sigma^2)T)) \\ &- qe^{-rT}\Phi(\frac{1}{\sigma\sqrt{T}}(\log(\frac{P_0}{q}) + (r - \frac{1}{2}\sigma^2)T)). \end{aligned}$$

□

Exercise 15.8.8. Consider the price formula (15.8.5), with r, σ, P_0 and q fixed positive quantities.
(a) What happens to the price (15.8.5) as $T \searrow 0$?Does this result make intuitive sense?

(b) What happens to the price (15.8.5) as $T \nearrow \infty$?Does this result make intuitive sense?

Solution.(a)

$$\lim_{T\to 0^+}(P_0\Phi(\frac{(\log(\frac{P_0}{q})+(r+\frac{1}{2}\sigma^2)T)}{\sigma\sqrt{T}})-qe^{-rT}\Phi(\frac{(\log(\frac{P_0}{q})+(r-\frac{1}{2}\sigma^2)T)}{\sigma\sqrt{T}}))$$

equals to $P_0 - q$ if $P_0 > q$, $\frac{1}{2}(P_0 - q)$ if $P_0 = q$, and 0 if $P_0 < q$.
This result means that a fair price for the option to purchase the stock at a fixed $T \simeq 0$, for a fixed price $q > 0$ and the current price P_0 of the stock is $P_0 - q$, $\frac{1}{2}(P_0 - q)$, or 0 provided that $P_0 > q, P_0 = q$ or $P_0 < q$, respectively.

(b)

$$\lim_{T\to\infty}(P_0\Phi(\frac{(\log(\frac{P_0}{q})+(r+\frac{1}{2}\sigma^2)T)}{\sigma\sqrt{T}})-qe^{-rT}\Phi(\frac{(\log(\frac{P_0}{q})+(r-\frac{1}{2}\sigma^2)T)}{\sigma\sqrt{T}}))=P_0.$$

This result means that a fair price for the option to purchase the stock at a fixed $T \simeq \infty$, for a fixed price $q > 0$ and the current price P_0 of the stock is P_0.$\square$

Appendix A

Mathematical Background

Exercise A.3.2. Use the definition of limit to prove that:
(a) $\lim_{n\to\infty} \frac{1}{n} = 0$.
(b) $\lim_{n\to\infty} \frac{1}{n^k} = 0$ for any $k > 0$.
(c) $\lim_{n\to\infty} 2^{\frac{1}{n}} = 1$.

Solution.(a) Let $1 > \epsilon > 0$ be given. Then, for $N = [\frac{1}{\epsilon}] + 1$ and any $n > N$ it follows that:

$$|\frac{1}{n} - 0| = \frac{1}{n} < \frac{1}{N} = \frac{1}{[\frac{1}{\epsilon}] + 1} < \frac{1}{\frac{1}{\epsilon}} = \epsilon.$$

(b) Let $1 > \epsilon > 0$ be given. Then, for $N = [(\frac{1}{\epsilon})^{1/k}] + 1$ and any $n > N$ it follows that:

$$|\frac{1}{n^k} - 0| = \frac{1}{n^k} < \frac{1}{N^k} = \frac{1}{([(\frac{1}{\epsilon})^{1/k}] + 1)^k} < \frac{1}{((\frac{1}{\epsilon})^{1/k})^k} = \epsilon.$$

(c) Note that :

$$2^{1/n} - 1 = \frac{1}{\sum_{k=0}^{n-1} 2^{k/n}} < \frac{1}{\sum_{k=0}^{n-1} 1} = \frac{1}{n} : \quad n \in \mathbb{N}.$$

Thus, by similar argument presented in part (a) the assertion follows. □

Exercise A.3.8. Prove that $\sum_{i=2}^{\infty}(1/i\log(i)) = \infty$.

Solution. The reader is reminded that by Cauchy condensation test, for a non-increasing sequence $f(n)$ of non-negative real numbers, the series $\sum\limits_{i=1}^{\infty} f(i)$ converges if and only if the "condensed" series $\sum\limits_{i=0}^{\infty} 2^i f(2^i)$ converges. Here, for $f(i) = \frac{1}{i\log(i)}$, the condensed series equals $\sum_{i=1}^{\infty} \frac{1}{i\log(2)}$, which equals infinity by statement $(A.3.7)$ with $a = 1$. The result follows. □

Exercise A.4.4. (a) Let $R, S \subset \mathbb{R}$, each be non-empty and bounded below. Prove that $\inf(R \cup S) = \min(\inf R, \inf S)$.

(b) Prove that this formulae continues to hold if $R = \phi$ and or $S = \phi$.
(c)State and prove a similar formulae for $\sup(R \cup S)$.

Solution. (a) Since $R \subset R \cup S$, it follows by definition that $\inf(R) \geq \inf(R \cup S)$. Similarly, $\inf(S) \geq \inf(R \cup S)$. Accordingly, $\min(\inf(R), \inf(S)) \geq \inf(R \cup S)$. The other side of inequality is a direct application of Proposition A.4.2 (Check !).

(b) This is strightforward result of $\inf \phi = 0$.

(c) Note that for any $A \subset \mathbb{R}$:

$$\begin{aligned} \inf(A) &= -\sup(-A), \quad (*) \\ \sup(A) &= -\inf(-A). \end{aligned}$$

Hence, by an application of part(a) and twice application of $(*)$:

$$\begin{aligned} \sup(R \cup S) &= -\inf(-(R \cup S)) \\ &= -\inf(-(R) \cup -(S)) \\ &= -\min(\inf(-R), \inf(-S)) \\ &= -\min(-\sup(R), -\sup(S)) \\ &= \max(\sup(R), \sup(S)). \quad \square \end{aligned}$$

Exercise A.4.6. Suppose $\{a_n\} \to a$ and $\{b_n\} \to b$, with $a < b$. Let $c_n = a_n$ for n odd and $c_n = b_n$ for n even. Compute $\liminf_n c_n$ and $\limsup_n c_n$.

Solution. By repeated applications of Exercise A.4.4 and continuty of bivariate function $f(x, y) = \min(x, y)$ on the real line we have:

$$\begin{aligned} \liminf_n c_n &= \lim_n \inf(\{c_k\}_{k=n}^{\infty}) \\ &= \lim_n \inf(\{c_{2k}\}_{k=[n/2]+1}^{\infty} \cup \{c_{2k+1}\}_{k=[n/2]+1}^{\infty}) \\ &= \lim_n \inf(\{b_{2k}\}_{k=[n/2]+1}^{\infty} \cup \{a_{2k+1}\}_{k=[n/2]+1}^{\infty}) \\ &= \lim_n \min(\inf\{b_{2k}\}_{k=[n/2]+1}^{\infty}, \inf\{a_{2k+1}\}_{k=[n/2]+1}^{\infty}) \\ &= \min(\lim_n \inf\{b_{2k}\}_{k=[n/2]+1}^{\infty}, \lim_n \inf\{a_{2k+1}\}_{k=[n/2]+1}^{\infty}) \\ &= \min(b, a) \\ &= a. \end{aligned}$$

Similarly (check!):

$$\limsup_n c_n = b. \quad \square$$